VOL. 13 No. 5, pp. 109-144 NOVEMBER 1951

Philips Technical Review

DEALING WITH TECHNICAL PROBLEMS
RELATING TO THE PRODUCTS, PROCESSES AND INVESTIGATIONS OF
THE PHILIPS INDUSTRIES

EDITED BY THE RESEARCH LABORATORY OF N.V. PHILIPS' GLOEILAMPENFABRIEKEN, EINDHOVEN, NETHERLANDS

A NEW HIGH-PRESSURE MERCURY LAMP WITH FLUORESCENT BULB

by J. L. OUWELTJES, W. ELENBAAS and K. R. LABBERTÉ.

621.327.312 :535.37

In the last decade the use of fluorescent substances has brought about a revolution in illumination engineering, the principal example of which was the low-pressure mercury lamp with fluorescent envelope (the "TL" lamp). Before the triumphal advent of this low-pressure lamp, however, the idea of applying fluorescent substances for the purposes of illumination had already materialized in the high-pressure mercury lamp. It is interesting to give thought to the differences in the problems encountered with these two kinds of lamps. Thanks to the development of new materials and methods, the high-pressure mercury lamp with fluorescent bulb has now also an opportunity to take an important place in the field of illumination, inter alia for industrial and street lighting.

The HP lamp

The principle of the electric discharge in mercury vapour of low pressure was applied several decades ago in tubes for illuminated advertising signs. For lighting purposes, however, this could not be applied because of the blue colour of the light and the very low efficiency of the light source in lumens per watt; the radiation produced by the discharge consists for about 80% of ultra-violet with a wavelength of 2537 Å, only about 5% being in the visible part of the spectrum.

About 1935 a lamp was made which was based on a discharge in mercury vapour under high pressure 1)) (e.g. 5 atm or higher). This meant a great step forward in the direction of applying such discharges for lighting purposes. The high-pressure mercury lamp (HP lamp 2))) has a radiation spectrum as represented in *fig. 1*. A large part of the energy of this lamp is also radiated in the ultra-violet part of the spectrum, the main wavelengths being 2537 Å and 3650 Å, but owing to the high pressure there is at the same time a continuous spectrum extending over the whole of the visible range; again due to the high pressure, the spectral lines in the visible range are also of a relatively greater intensity. The result is that quite a reasonable efficiency is obtained, viz. 35 to 40 lumens/watt. It is true that the colour of this light is still rather bluish and the colour rendering leaves much to be desired, but the high efficiency compared with the incandescent lamp made the use of these high-pressure mercury lamps so attractive that they soon found application in various cases, sometimes in combination with incandescent lamps.

Fig. 2 shows the construction of the HP lamp in its present form. The discharge takes place in a small tube of fused silica containing a certain amount of mercury together with a rare gas under a pressure of a few cm Hg. This rare gas and the provision of an auxiliary electrode enables the lamp to be started with normal mains voltage (220 V$_{\sim}$). The heat generated by the discharge causes the mercury to evaporate and the pressure in the tube rises to a few atm (according to the type of tube). The discharge tube is mounted in a bulb of the same shape as that of an incandescent lamp

---

1) W. Elenbaas, Discharges in mercury vapour under high pressure (in Dutch), De Ingenieur **50**, E83-90, 1935; C. Bol, A new mercury lamp (in Dutch), De Ingenieur **50**, E91-92, 1935: W. de Groot The emission and absorption spectra of mercury vapour at very high pressures (up to 300 atm) (in Dutch), De Ingenieur **50**, E92-94, 1935. See also G. Heller, The mercury vapour lamp HP 300, Philips Techn. Rev. **1**, 129-134, 1936.

2) Sometimes the HP lamps, together with the SP lamps, which are water-cooled and in which pressures of about 100 atm occur, are denoted as "super-high-pressure" mercury lamps, the name of "high-pressure" lamp then being reserved for lamps with a mercury pressure of about 1 atm (HO lamps).

and filled with nitrogen. The gas jacket thus formed round the discharge tube has a two-fold function. In the first place it protects the current leads fused in the silica, which are heated to a very high temperature, so that they would rapidly oxidize at the end where they come into contact with air. In the second place the nitrogen affords a certain thermal insulation: if the tube were exposed to the open air the thermal dissipation through convection would be so great that upon the current being switched on the tube would only very slowly heat up and the mercury would take a long time to evaporate completely, if at all. (A more complete insulation, by evacuating the bulb — in which case the only dissipation of heat would be by radiation —, is not desirable because then, in order to avoid the wall of the tube reaching too high a temperature, only a smaller power can be allowed to develop in the discharge tube.)

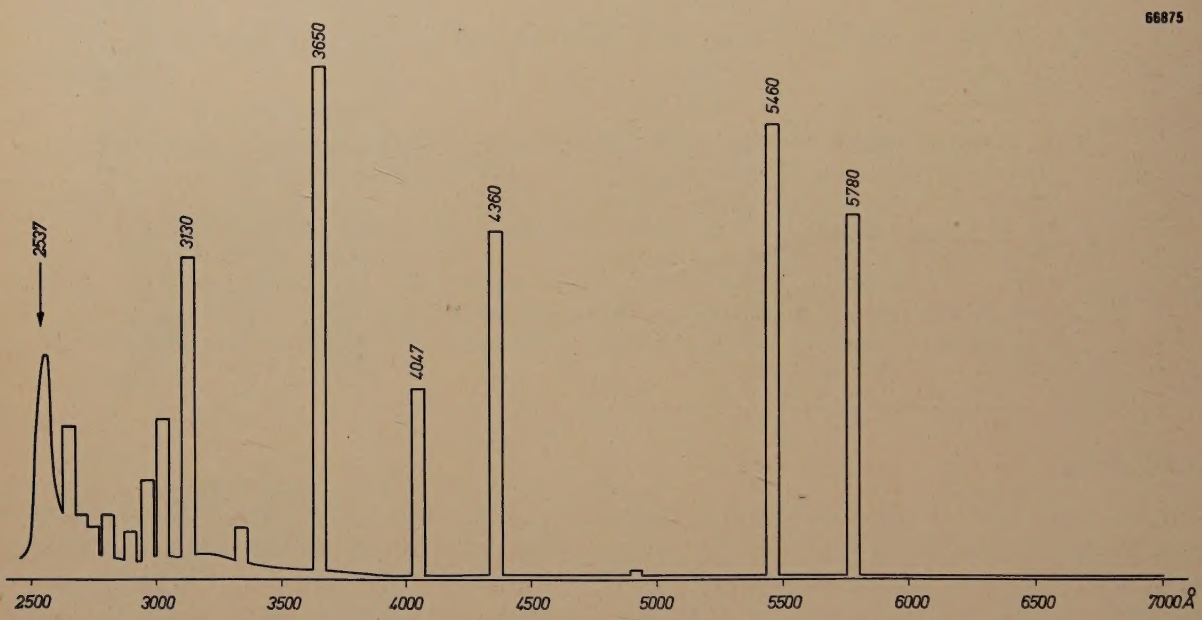


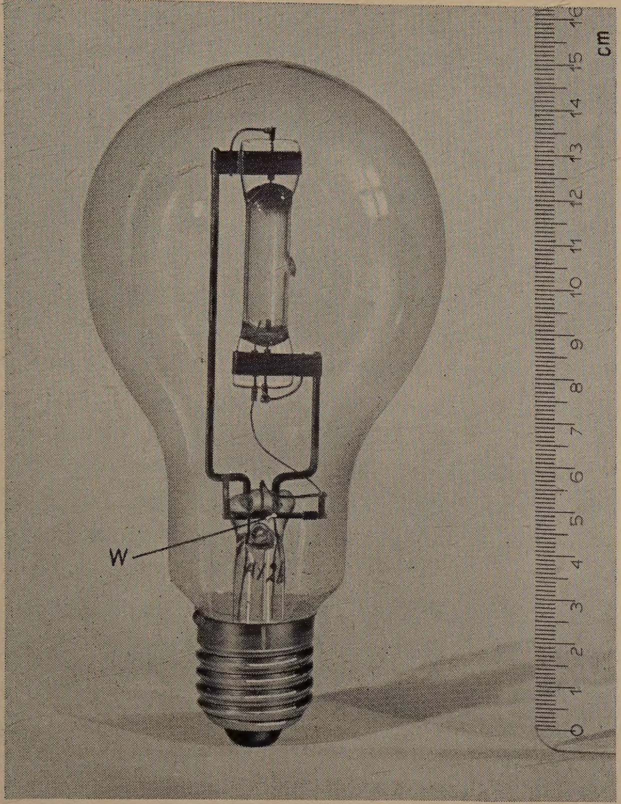
Fig. 1. The spectral energy distribution of the HP lamp (without glass bulb). The spectral lines superposed on the continuum have all been drawn relatively too low and with the same exaggerated width, such that the area covered by each line represents the energy emitted therein. About 1½ times as much energy is radiated in the ultra-violet as in the visible part of the spectrum.

Some other details of the construction are explained in the subscript to fig. 2. We would only mention here — because they will be referred to again later — the strips of molybdenum, serving to carry the current through the fused silica, and the inside frosting of the bulb which hides the inside of the lamp from view and gives the lamp a pleasant appearance when it is not burning while at the same time providing for a not too great surface brightness when the lamp is alight.

Possibilities for improving the colour rendering of the HP lamp

Considering the high efficiency and other good properties of the HP lamp, there was every inducement to investigate whether the colour rendering could not be so improved as to make these lamps suitable for more universal use. The lack of colour rendering is most noticeable with objects the colour of which is predominantly red: under the HP lamp such objects assume a brownish hue. This can be understood from a look at the spectrum in fig. 1. In the red, i.e. roughly between 6000 Å and 7000 Å, the mercury vapour has no spectral lines of any intensity worth mentioning, and also the continuous spectrum of the HP lamp is weak in that range. The radiation of the HP lamp therefore needs supplementing in the red.

Attempts have been made to attain this by various means: (1) By adding a little cadmium to the mercury vapour. It is true that the cadmium vapour does to a certain extent yield the desired contribution in the red, but this artifice is accompanied by a considerable lowering of the efficiency and a shortening of the life of the lamp. (2) The choke (or low power factor transformer), usually connected in series with the mercury discharge in order to stabilize it, was replaced by a filament mounted together with the discharge tube in one bulb. The incandescent light from the filament, being very rich in the red, supplemented



66793

Fig. 2. The HP lamp. Inside, in the centre of the bulb, is the tube of fused silica in which the discharge takes place in mercury vapour of high pressure. The current is fed to the electrodes via leading-in strips of molybdenum (almost concealed by the supports of the discharge tube) sealed airtight in the tube. Beside the lower main electrode is an auxiliary electrode connected to the other main electrode via a current-limiting resistor *W* and serving to facilitate ignition (ignition voltage about 180 V). Usually the bulb is frosted on the inside (here it has been left clear to show the mounting inside). It is filled with nitrogen under a pressure of 50 cm Hg (in the cold state).

the radiation from the mercury discharge so as to give it a white light with good colour rendering. The contribution of light needed from the filament, however, is of such a high percentage that the efficiency of these blended-light lamps, named ML lamps, is greatly inferior (about 20 lm/W) to that of the HP lamp. (3) The most elegant solution is that whereby the otherwise useless ultra-violet radiation from the mercury vapour is converted by means of a fluorescent substance into the red radiation desired. In this way the efficiency in lumens/watt need not by any means suffer but may in principle even be made better than that of the HP lamp. This line of thought led to the construction of the HPL lamp, which is a high-pressure mercury lamp having a bulb coated on the inside with a fluorescent layer.

Before dealing more closely with this idea and the manner in which it was materialized, further development will be briefly outlined.

The first HPL lamps brought onto the market about 1937 were far from being ideal as regards colour rendering. In those days, however, when the art of producing fluorescent substances (phosphors) was still more or less in its infancy, we had to consider ourselves fortunate in having found a phosphor which could be excited by the radiation of both 2537 Å and 3650 Å and which, moreover, answered fairly well the other requirements set (about which more will be said later). It is to be realized that once thoughts had turned to the use of fluorescent substances it was only obvious to try to make use of them also for the low-pressure mercury lamp for converting the lost radiation of 2537 Å into visible light. Such could not be done at that moment because no phosphors were then known which satisfied the entirely different requirements for that lamp.

As is known, some time later however phosphors were also found which were suitable for the low-pressure lamp, and thus the tubular fluorescent lamps of the "TL" type made their appearance. These very soon became enormously popular, so much so in fact that the previously developed but still immature HPL lamp was relegated to the background. However, investigations with the HPL type of lamp were continued, though for a long time without any great success. It was not until in the course of last year that a phosphor was discovered which has made a decided improvement of the HPL lamp possible. The new type of HPL lamp that is now in production promises to take up a field of application of its own, side by side with the so popular "TL" lamp.

What is required of a phosphor for the HPL lamp and what is to be expected from it

As already indicated, to improve the colour rendering of the HP lamp a phosphor is needed which is excited by radiation of wavelengths in the range of 2537 and 3650 Å and which gives an emission spectrum lying mainly in the red.

Further, the phosphor has to satisfy the following requirements:

1) It must not absorb the visible light of the mercury discharge to any appreciable extent, so as not to affect the efficiency of the lamp. A strong selective absorption is also undesired; apart from the coloured appearance it would give the lamp while it is not burning, which effect in some cases

---

[3] E. L. J. Matthews, The blended-light lamp and other mercury lamps with improved colour rendering, Philips Techn. Rev. 5, 341-347, 1940. J. Funke and P. J. Oranje, The development of blended-light lamps, Philips Techn. Rev. 7, 34-40, 1942.

may be unpleasant, selective absorption might again spoil the colour rendering.

2) Although any phosphor has the tendency to lose its fluorescent power as the temperature increases, this phosphor must still give a reasonably strong fluorescence at a temperature of 150 °C to 200 °C, this being approximately the temperature reached by the bulb that is to be lined with the fluorescent layer. Of course the temperature of the bulb can be reduced somewhat by making the bulb larger, and in fact a step has been taken in that direction (the diameter of the old 120 W HPL 500 lamp was 130 mm, as compared with the 90 mm of the HP 500), but there is a limit to this because the fittings in which the HPL lamps are to be used should not be too large.

3) Finally there is the very important requirement that the phosphor must not be subject to dissociation under the strong ultra-violet irradiation (the photo-chemical action of which is further promoted by the high temperature). Anticipating what is stated farther on in this article, it may be said here that for many substances which otherwise appeared to be suitable for our purpose this requirement proved to be the stumbling block: the photochemical dissociation occurring with the substances first becomes noticeable in a reduction of the content of red light in the radiation from the lamp; later the fluorescent layer often turns grey, thus absorbing visible light and causing the total luminous flux of the lamp to diminish considerably.

Supposing that a phosphor were found which ideally answers these requirements, the question is what result is to be expected from it as regards the colour rendering of the lamp. This can be calculated in a simple manner. The colour rendering of a lamp is most easily judged with the aid of the "block diagram" devised by Bouma [4]). According to this method the spectral distribution of a light source is described by quoting the percentage of the total luminous flux supplied in each of the eight adjacent spectral "blocks" covering together the whole of the visible spectrum.

In *table I* figures are given for the HP lamp and for daylight; it should, however, be noted that these figures represent energy contributions and not contributions of luminous flux. The qualitative differences between the sources of light as considered in each block are, of course, the same in both cases, and by quoting the energies it is possible to give in the same table also the contributions in two ultra-violet blocks. Now these ultra-violet energy contributions of the HP lamp are to be imagined as being transferred to the red blocks 6100 Å - 6600 Å and 6600 Å - 7200 Å, where the radiation from the HP lamp shows such a marked deficit. In this transference account has to be taken of the depreciation in energy taking place in any process of fluorescence: for one quantum of ultra-violet energy $h\nu_1$ (given a quantum efficiency of 100%) at most one energy quantum $h\nu_2$ of the changed wavelength can be obtained, and since a red quantum contains only about half the energy of one of our ultra-violet quanta ($\nu_2 \approx \frac{1}{2}\nu_1$) in the transference the energy has to be halved. In each of the two red blocks there is then a relative energy contribution of about 35. In reality, of course, a quantum efficiency of 100% cannot be reached, and, moreover, to ensure sufficient transmission of the visible light from the mercury discharge, the phosphor can only be applied in such a thin layer that the ultra-violet radiation is not completely absorbed. The conclusion, therefore, is that in the most favourable case possible of realization the colour rendering of daylight in the red can just about be reached [5]). The improvement that can be attained appears to be indeed quite sufficient for the red colour rendering, provided — and this is the lesson

Table I. Block-diagram distribution of the spectrum of the HP lamp and of average daylight. Contrary to the usual custom the relative energy contributions in the blocks are given (slightly rounded off), instead of the relative luminous flux contributions. Also the ultra-violet energy in two spectral sections is given, in the same relative measure as the visible radiation.

	< 3000 Å	3000-4000	4000 No. 1	4200 2	4400 3	4600 4	5100 5	5600 6	6100 7	6600 8 7200
HP 125 W	60	80	13	24	1	2	31	26	1,5	1,5
Average daylight	0	9	6	6	7	18	17	16	14	16

---

[4]) P. J. Bouma, Colour reproduction in the use of different sources of "white" light, Philips Techn. Rev. 2, 1-7, 1937. See also P. M. van Alphen, A photometer for the investigation of the colour rendering of various light sources, Philips Techn. Rev. 4, 66-72, 1939.

[5]) For this argument we are indebted to Dr. J. Voogd, and for the necessary connections between energy measurements in the ultra-violet and in the visible spectrum we have to thank J. Riemens.

to be learned from the foregoing — the fluorescence is utilized as far as possible to augment the red content of the radiation.

Phosphors tried out for the HPL lamp

The phosphor which until recently had always been used in HPL lamps is zinc cadmium sulphide, activated with copper, (Zn, Cd)S-Cu. The spectral energy distribution of the emission from this phosphor is represented in *fig. 3*, from which it is seen that there is a good maximum in the orange-red. Further, the "excitation spectrum" is not unfavourable; the phosphor is highly absorbent for all the ultra-violet radiation that is to be considered and it fluoresces with a reasonable quantum efficiency: upon activation with 3650 Å it is about 50% and with 2537 Å about 25%, measured at room temperature. At the working temperature (about 175 °C) these values are only slightly lower. The phosphor is also sufficiently durable during the life of the lamp. A serious drawback, however, is that this zinc-cadmium-sulphide phosphor is highly absorbent also for the blue radiation (an extension of the ultra-violet absorption band, which in itself is indispensable). Instead of a lack of red the light from the burning lamp now has a deficit in blue and is greenish in hue. Consequently under this old HPL lamp all practically white objects assume an unpleasant appearance, and this applies in particular to the human skin.

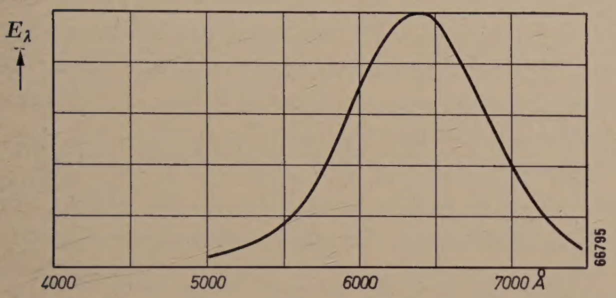


Fig. 3. Spectral energy distribution (in arbitrary units) of the emission of zinc cadmium sulphide activated with copper.

Various means have been tried in an attempt to remedy this failing of the zinc-cadmium-sulphide phosphor. Attemps were made to reduce the absorption of blue by applying the phosphor to the bulb in the form of a finer powder, but this appeared to be impossible without deteriorating the phosphor. For a long time the zinc-cadmium-sulphide phosphor used in the old HPL lamps was also replaced partly by a blue-fluorescing phosphor (ZnS-Ag), but this solution was not entirely satisfactory and larger bulbs had to be used on account of the admixture being rather sensitive to temperature.

Thus right from the beginning there were reasons for looking for entirely different fluorescent substances suitable for correcting the colour of the light emitted by the HP lamp. This became all the more imperative with the advent of the "TL" lamp, not only because this lamp with its high efficiency and excellent colour rendering made it necessary to lay down more stringent standards, but also because it had thereby been learned that phosphors exist with a quantum efficiency up to about 80%, as compared with the 50% and 25% for the zinc-cadmium-sulphide phosphor.

Of the various phosphors which subsequently became available and which were tried out for use in the HPL lamp four will be discussed here, viz. cadmium borate, calcium phosphate, the group of halophosphates and a group of phosphors with quadrivalent manganese as activator.

Cadmium borate, with divalent manganese as activator, absorbs only the short-wave ultra-violet; thus the emission near 3650 Å is unused. The emission spectrum, however, is very favourable, as is shown by curve *a* in *fig. 4*, and there is no absorption of visible rays, so that with this phosphor quite acceptable lamps could be made, be it with a somewhat larger bulb because the luminescence is considerably weaker even at 150 °C. Unfortunately there was such a great decline in luminescence during the useful life of the lamp that this phosphor could not be considered for practical use.

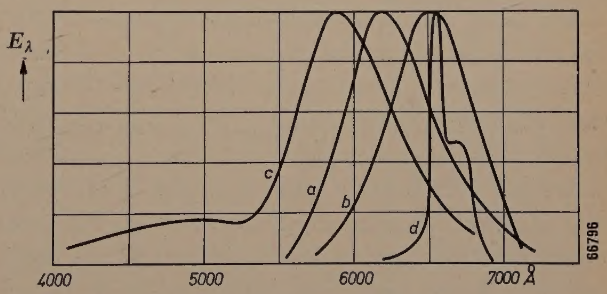


Fig. 4. Spectral energy distribution of four phosphors tried out for their suitability for use in the HPL lamp (maxima all taken to be equally high):
a) Cadmium borate activated with divalent manganese.
b) Calcium phosphate activated with cerium and manganese.
c) Halophosphates activated with antimonium and manganese.
d) Magnesium titanate activated with quadrivalent manganese.

Much the same applies also to calcium phosphate (activated by cerium and manganese). This substance absorbs only the short-wave ultra-violet; in daylight it is white and it shows a beautiful red fluorescence (fig. 4*b*) with a fairly good quantum efficiency

(50%). In regard to temperature-sensitivity this phosphor is even better than cadmium borate, the efficiency remaining absolutely constant up to about 300 °C. But here again no remedy could be found for the rather rapid decline in luminescence during the useful life of the lamp.

The third kind of phosphors mentioned, the halophosphates, are of the general formula $3Ca_3(PO_4)_2CaX_2$, where X stands for chlorine or fluorine. In this case the activators are antimony and manganese. These phosphors have gained much in practical importance through their application in "TL" lamps. Essential for this application is a strong luminescence upon excitation by 2537 Å and sufficient resistance against the dissociating effect of these rays, as also against the effect of the mercury vapour, with which the fluorescent layer in "TL" lamps comes into contact. The halophosphates possess the above-mentioned properties and also retain their fluorescence at high temperatures (this is a property that is not really essential for "TL" lamps because the wall of the "TL" lamp does not reach a temperature higher than 40 °C to 50 °C); further, they are also resistant against dissociation at the high temperatures prevailing in an HPL lamp. They do not, however, answer the specific emission requirements for the HPL lamp: at the best only a spectral distribution can be obtained as drawn in fig. 4*c*, with a maximum at about 5900 Å. This is not sufficient for improving the red content of the radiation from the HP lamp.

Finally there are the substances containing quadrivalent manganese as activator, several of which have been throughly investigated in the Philips Laboratory at Eindhoven 6). All these show a deep red fluorescence. Magnesium titanate, which belongs to this group of phosphors, appeared to offer the best prospects as regards suitability under the conditions occurring in the HPL lamp. The emission spectrum is given in fig. 4*d*. The excitation spectrum, unfortunately, is less satisfactory, for, though the rays with wavelength 2537 Å are well absorbed, the quantum efficiency of the conversion of those short-wave rays is practically nil. Attempts were made to remedy this evil by causing the ultra-violet HP radiation to fall first upon a layer of some other suitable phosphor (e.g. cadmium borate) which absorbs only the short-wave rays and converts them into red. However this idea has not been put into practical execution because magnesium titanate also has the drawback of its fluorescence being too sensitive to temperature: above 100 °C it diminishes rapidly.

Magnesium arsenate

Although magnesium titanate, belonging to the group of Mn^{4+} phosphors, did not answer the requirements, it was this very same group which yielded the phosphor with which the problem of the HPL lamp was ultimately solved, viz. magnesium arsenate 7).

Having already become familiar with the problems of preparing Mn^{4+} phosphors and with the characteristics of these substances, it was very soon possible, once attention had been drawn to it, to prepare this new phosphor in a form suitable for application to lighting technique.

The properties of this magnesium arsenate 8) proved to be highly favourable for its application in the HPL lamp. Its fluorescence is bright red.

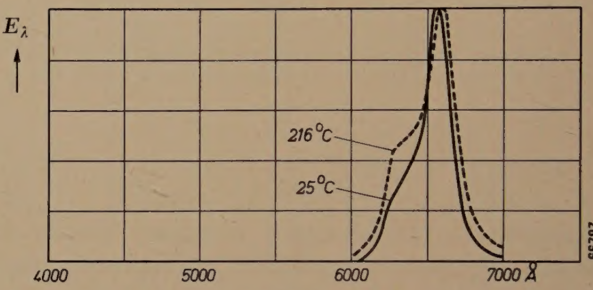


Fig. 5. Spectral energy distribution of magnesium arsenate activated with quadrivalent manganese, at 25 °C and at 216 °C (maxima taken to be equally high).

The spectral distribution of the emission (*fig. 5*) shows a steep peak at 6560 Å. This phosphor absorbs the whole of the ultra-violet radiation spectrum of the HPL lamp and has a quantum efficiency of about 75 % when excited either with 2537 Å or with 3650 Å (and also with intermediate wavelengths). There is a slight absorption of blue light — considering the lack of any sharp demarcation of the absorption bands of solids this absorption of some blue is inevitable if the whole of the long-wave ultra-violet is to be absorbed — but the resultant faintly yellow colour of the magnesium arsenate in daylight is not troublesome, and, contrary to the case with the old zinc-cadmium-sulphide lamps, there is no question of any green coloration of the resultant light from the HPL lamps in which this new phosphor is used. Another very favourable property of magnesium arsenate

---

6) F. A. Kröger, Some Aspects of the Luminescence of Solids, Elsevier Publ. Co., Amsterdam, 1948, in particular pp 57-106.

7) M. Travnicek of Graz (Austria) brought this substance to our notice.

8) The investigations into the properties of the new phosphor, as also the first investigation into the conditions for obtaining a high efficiency from it, were carried out in the Philips Laboratory by Dr. F. A. Kröger.

is the fact that it shows very little sensitivity to temperature, the quantum efficiency still being about 65% at 300 °C (*fig. 6*).

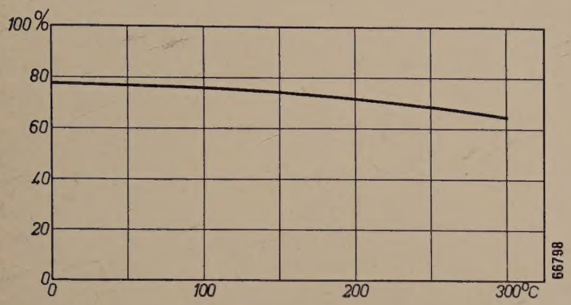


Fig. 6. The quantum efficiency of the magnesium arsenate phosphor upon excitation with 2537 Å, as a function of the temperature.

Chemical behaviour of magnesium arsenate in an HPL lamp

After this summing up of the excellent properties of magnesium arsenate it seems like an anti-climax to have to state that in the beginning this phosphor, like so many others, failed to satisfy the last requirement, namely that of durability. The red content of the light from the first HPL lamps made with magnesium arsenate declined rapidly.

Contrary, however, to the case with other phosphors which came to grief over this stumbling block, a simple means was found of improving the durability of magnesium arsenate. All that was necessary was to replace the nitrogen filling of the bulb of the HPL lamp by carbon dioxide. It is a most remarkable phenomenon that then during the first 20 to 50 burning hours the red content of the light increases, sometimes by a factor of as much as $1^1/_2$, after which it gradually decreases, but after 2000 burning hours it is still about 80% of the peak value.

Though the mechanism by which CO_2 improves the resistance of the magnesium-arsenate phosphor has not yet been fully explained, it was by fairly logical reasoning that the choice of CO_2 came to be made.

Magnesium is one of the basest metals (i.e. difficult to reduce). It was therefore probable that the decline in the fluorescence of the phosphor is not due to any separation of metal (in this case Mg), as was found to be the case with the other phosphors, but that it is to be ascribed to a reduction to a lower valency of the small amount of Mn^{4+} present (for the best yield of light a composition something like 10 $MgO \cdot As_2O_5 \cdot 0.01$ Mn is chosen). Chemical analysis confirmed the supposition that the decline in fluorescence of the phosphor is accompanied by a reduction from Mn^{4+} to Mn^{2+}. Experiments were therefore made with a small addition of oxygen to the nitrogen filling of the bulb of the lamp. It then appeared that (at a bulb temperature of 150 °C to 200 °C) with the addition of 0.3 % to 0.5 % of O_2 and under a nitrogen pressure of 50 cm Hg there was indeed no longer any decline at all in the red content of the light.

However, with such an admixture of oxygen another difficulty arose, viz.: the extremely thin leading-in strips of molybdenum which are fused airtight in the silica for carrying the current through the "pinches" at the ends of the mercury-discharge tube (*fig. 7a*; see also fig. 2) tend to oxidize in an atmosphere containing O_2, thereby expanding and after some time causing the pinches to crack. The higher the temperature at the outward extremity of the molybdenum strip, the more likely is this to occur. With the dimensions of the pinches usual in HP lamps this temperature amounts to about 500 °C, and with a partial oxygen pressure of 0.15 cm Hg (see above) the pinch cracks even after the lamp has been switched on for some tens of hours only.

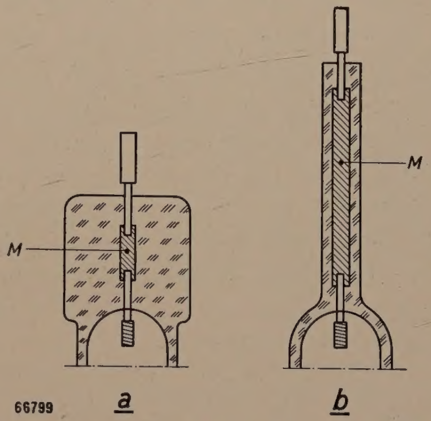


Fig. 7. Construction of the pinch of the discharge tube in the HP and HPL lamps. When the lead is being fused in, it has to be heated up to the softening temperature of the fused silica, i.e. 1600 °C to 1700 °C. Of the few metals capable of withstanding this temperature molybdenum is the most suitable. Since, however, the thermal expansion coefficient of molybdenum is much greater than that of silica, instead of a leading-in wire, use is made of a strip of molybdenum *M*, rolled out extremely thin (13 μ thick) ,which, being readily deformable, adheres to the fused silica in spite of the differences in expansion. *a*) Pinch of normal length (drawn to approximately true size). *b*) The pinch made longer so that the outward extremity of the leading-in strip does not become so hot as in (*a*), thus reducing the rate of oxidation of the molydenum when oxygen is added to the nitrogen filling of the bulb (the round leads are not sealed airtight in the fused silica and hence oxygen can penetrate from the outside of the tube and reach the molybdenum).

This can be remedied by making the pinch longer (fig. 7*b*). Of course oxidation then still takes place, but at a slower rate because the temperature at the danger point is lower. In this way the durability of the lamp is no longer determined by that of the pinch construction. HPL lamps made in this manner were quite satisfactory, but their manufacture proved to be too costly because the longer pinches could not be made mechanically; furthermore, the O_2 percentage of the filling is very critical, because if it is slightly too low the yield of light from the phosphor diminishes rapidly and if it is a little on the high side then the leading-in strips oxidize too quickly.

It was then thought that perhaps in itself a partial oxygen pressure much lower than 0.15 cm Hg would suffice to maintain the light yield of the phosphor, and that the "high" filling pressure of 0.15 cm Hg, which proves to be fatal for the molybdenum strips, is only needed on account of the fact that the various metal parts in the lamp selectively "catch" the oxygen and more or less retain it ("getter" action). In that case matters might be improved by doing without the oxygen in the filling of the bulb and using instead an inert gas which, under the influence of the high temperature and possibly also under the influence of the ultra-violet radiation, continuously gives off a little oxygen, just sufficient to prevent the reduction of Mn^{4+}, and too little to cause any serious oxidation of the molybdenum strips. CO_2 seemed to be suitable for this purpose and, indeed, it yielded the excellent results already mentioned.

Confirmation of the line of reasoning followed is found in the fact that the desired effect is obtained even with only a very low CO_2 pressure, viz. about 0.5 cm Hg. In the case of dissociation of the carbon dioxide to a degree of say 5%, which is plausible, according to the formula

$$2\ CO_2 \rightarrow 2\ CO + O_2,$$

the partial oxygen pressure is then no more than about 0.01 cm Hg.

With an oxygen filling the metal parts can take up only a certain amount of O_2. The larger the volume of the bulb, the smaller will be the resultant reduction of the O_2 pressure, and it may therefore be expected that then a lower initial pressure suffices in order to keep the O_2 pressure above the minimum required for the phosphor during the desired lifetime. This expectation was fully confirmed by a test whereby the O_2 volume was increased by connecting the bulb of the HPL lamp to a large reservoir via a small tube.

It appears that the CO_2 pressure is not at all critical. One can safely increase the pressure to 50 cm Hg. such as is usual for the nitrogen filling of the HP lamps. The fact that the molybdenum strips do not then oxidize too quickly is apparently due to the partial O_2 pressure not rising proportionately with the higher CO_2 pressure: the degree of dissociation of the carbon dioxide decreases with the increasing pressure, as follows qualitatively directly from the reaction equation given above (the volume on the right is $1^1/_2$ times that on the left).

Construction and properties of the new HPL lamps

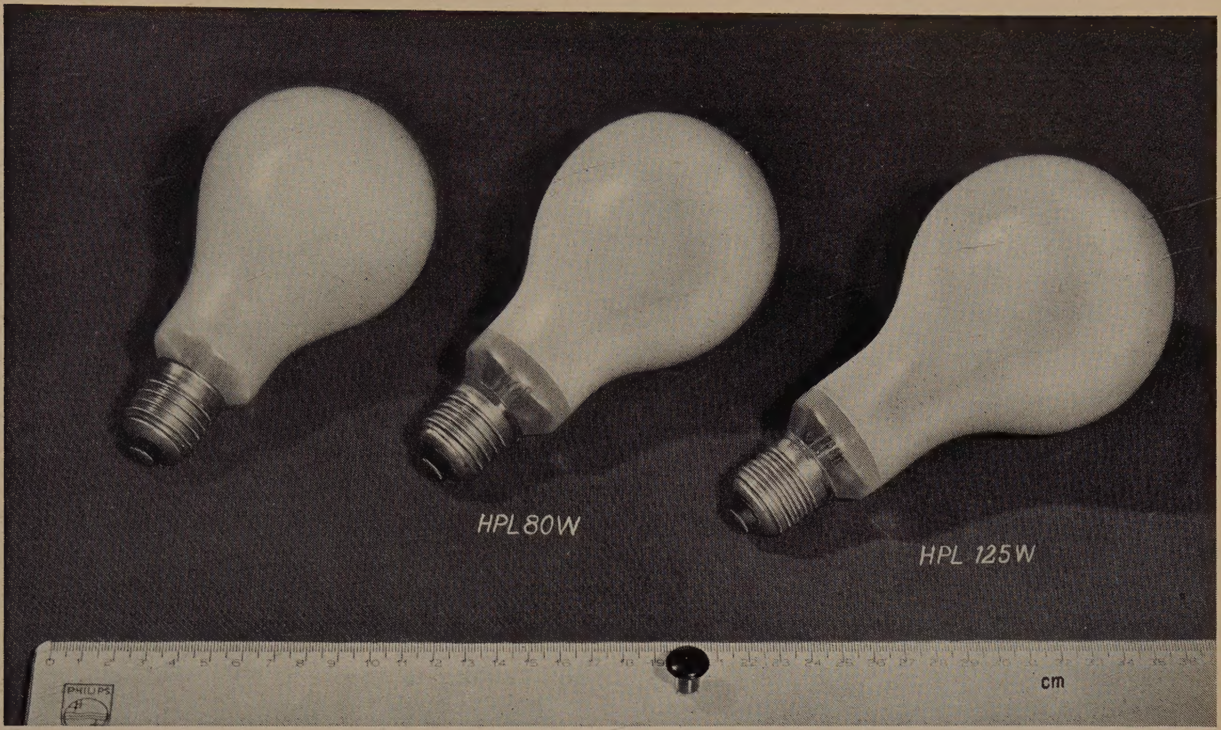
HPL lamps are now being made in two wattages, viz. 80 W and 125 W. *Fig. 8* shows these two lamps and, for comparison, a 150 W "Argenta" incandescent lamp. *Table II* gives the principal data for these lamps, for an HP lamp and for daylight.

As follows from the foregoing general considerations, in aiming at the maximum possible increase of the red content of the light emitted by the HP lamp one's efforts are limited by the fact that for complete absorption of the ultra-violet such thick layers of phosphor would have to be applied that the visible radiation from the mercury discharge would be reduced too much: notwithstanding the gain in the red, the total yield of light from the lamp would then be reduced. Now the thickness of the phosphor layer has been so chosen that the gain in the red and the loss over the entire visible spectrum just about compensate each other; in other words, the efficiency of the new HPL lamp is practically equal to that of the HP lamp.

The red content of the light given in the table and serving as a rough measure for the improvement reached in the colour rendering is the sum percentage of the contributions to the luminous flux in the seventh and eighth blocks. For a better

Table II. Properties of the new HPL lamp, compared with those of an HP lamp, an incandescent lamp and daylight.

		HPL 80 W	HPL 125 W	HP 125 W	150 W incand. lamp (2850 °K)	Daylight
Luminous flux (lm)		3000	5000	5000	2000	—
Efficiency (lm/W)		37,5	40	40	13	—
Percentage of red		8,8	7,5	1,3	18	11
Colour point in the colour triangle.	$x =$	0,390	0,385	0,319	0,448	0,320
	$y =$	0,434	0,425	0,405	0,407	0,330
Luminance (cd/cm²)		5	8	25	3	—
Diameter of bulb (mm)		80	90	90	80	—



66794

Fig. 8. The new HPL 80 W and HPL 125 W lamps coated on the inside with magnesium-arsenate phosphor. When not alight their colour is practically white. For comparison an "Argenta" 150 W incandescent lamp is also shown.

judgment of the colour rendering *table III* gives the complete block diagram distribution for the radiation of the HPL 80 W and the HPL 125 W, together with the distributions for the HP lamp and for daylight as already given in table I (but there expressed in luminous flux percentages) and also those for the incandescent lamp and for a "TL" lamp ("white" type).

Table III. Block distribution (relative luminous flux contributions) of the spectrum of the new HPL lamp, compared with that of the HP lamp, daylight, an incandescent lamp and a "TL" lamp ("white" type).

	4000 4200 1	4200 4400 2	4400 4600 3	4600 5100 4	5100 5600 5	5600 6100 6	6100 6600 7	6600 7200 8
HPL 80 W	0,008	0,37	0,025	0,62	49,8	40,5	8,0	0,76
HPL 125 W	0,008	0,42	0,04	0,54	47,0	44,4	7,0	0,65
HP 125 W	0,02	0,76	0,14	1,66	55,1	41,0	1,2	0,09
Average daylight	0,027	0,23	0,80	10,8	40,8	36,2	10,3	0,73
150W incand. lamp (2850°K)	0,006	0,060	0,25	5,47	33,4	42,6	16,7	1,57
"TL" 40 W ("white") . .	0,010	0,35	0,30	4,80	38,0	43,0	13,0	0,43

An important field of application for the HPL lamp is the lighting of streets and squares in towns, With the high efficiency of these lamps a high level of illumination can be reached in a very economical way and the colour rendering is such that there is no longer any question of the human skin assuming an unpleasant hue under this light. Further, with the aid of inconspicuous fittings the light from the HPL lamp can be just as easily concentrated and directed as that from a frosted or opalised incandescent lamp. For this reason in a number of cases the HPL lamp may be preferred to the "TL" lamp, the application of which for street lighting is now being tried out in various places.

In this connection something has to be said about the ballasts required for the HP and HPL lamps.

Lamps which are to be installed separately are usually (when a mains voltage of 220 V$\sim$ is available) connected in series with a choke (*fig. 9a*). Upon the lamp being switched on, thus while it is in the cold state, a current flows through it which is about $1^1/_2$ times the normal working current. This starting current is strong enough — with the usual nitrogen or carbon dioxide filling of the bulb under a pressure of about 50 cm Hg — for the discharge tube to be gradually heated and the discharge to reach full strength in 3 to 4 minutes, during which the voltage across the lamp rises and the current drops. The power factor (cos φ) of the lamp with choke is about 0.5.

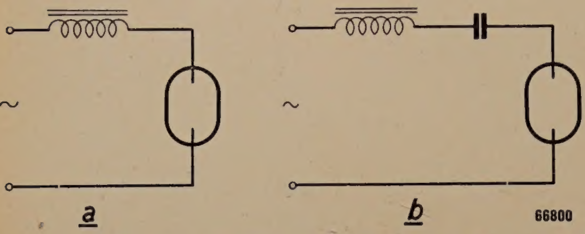


Fig. 9. As ballast for the HP and HPL lamps either a choke (*a*) or a combination of a choke and a capacitor (*b*) can be used. In the latter case the current is about just as much advanced in phase with respect to the voltage as it is lagging in the case (*a*). The starting current immediately after switching on the lamp is much lower in case (*b*) than in case (*a*).

This low power factor is no objection for single lamps, but in street lighting often whole rows of lamps have to be connected to one branch of the distributing network and in such a case a higher cos φ value is desired. This can be obtained by connecting half the number of lamps to a different ballast, viz. a combination of a choke and a capacitor (fig. 9*b*). The current is then about just as much advanced in phase with respect to the voltage as it is lagging in the circuit of fig. 9*a*. This brings the power factor for the whole installation to about 1. In the circuit of fig. 9*b* however the starting current is no stronger than the working current, and therefore, in order to ensure that the lamp will burn at full strength within a reasonable time after switching on, allowing also for severe frost, the discharge tube has to be thermally more insulated, i.e. the pressure of the bulb filling has to be made lower; for such a case a carbon dioxide pressure of about 20 cm Hg is quite suitable.

For this application of the HPL lamp it is therefore very convenient that the pressure of the carbon dioxide is not at all critical for maintaining the luminescence of the magnesium arsenate phosphor.

Summary. The high-pressure mercury lamp (HP) has a very good specific light yield, viz. 35 to 40 lm/W, but the colour rendering under the light of this lamp is poor, mainly due to a deficit in the red. The colour rendering can be greatly improved by converting the useless ultra-violet radiation from the mercury discharge into red light by means of a fluorescent substance. For the HPL lamp developed on this principle the fluorescent substance first used was zinc cadmium sulphide activated with copper. Upon excitation with the ultra-violet rays this substance shows an orange-red fluorescence with a reasonable quantum efficiency, it retains its fluorescence also at the temperature of about 150 °C to which it is exposed when applied to the inside of the bulb of an HP lamp, and while the lamp is alight this substance undergoes only a very gradual dissociation. A drawback however is that it absorbs the blue, so that the resultant colour rendering is nevertheless unsatisfactory. After a number of unsuccessful attempts to find more suitable phosphors, a very good solution has now been reached by employing magnesium arsenate activated with quadrivalent manganese. The quantum efficiency of this bright-red fluorescing phosphor is even appreciably higher than that of the phosphor first mentioned (viz. 75%, upon excitation either with 2537 Å or with 3650 Å) and less sensitive to temperature, whilst the absorption of blue is so slight as to be practically of no effect, so that the colour rendering is quite satisfactory. At first this phosphor was not sufficiently resistant to dissociation, but this has been remedied and excellent resistance obtained by filling the bulb of the HPL lamp with carbon dioxide instead of the nitrogen commonly used for HP lamps. The action of the carbon dioxide in this respect is to be ascribed to a little oxygen being given off which prevents a reduction from quadrivalent to divalent manganese in the phosphor. This effect can also be obtained by the direct admixture of a little oxygen to the normal nitrogen filling, but then there is a prohibitively rapid oxidation of the molybdenum leading-in strips of the discharge tube. The carbon dioxide pressure can be varied between 0.5 and 50 cm Hg, thus making it possible to adjust the thermal insulation of the discharge tube according to the starting current available. This possibility is of practical importance for the application of HPL lamps for street lighting.

THE IMAGE ICONOSCOPE, A CAMERA TUBE FOR TELEVISION

by P. SCHAGEN, H. BRUINING and J. C. FRANCKEN.

621.385.832:621.397.611

The oldest television camera tube, the iconoscope, is now used only for transmitting still pictures (e.g. the signal picture of a certain station) and film pictures. Further development of camera tubes in Europe has followed a course different from that in America. In the U.S.A. the image orthicon has become predominant, whilst in Europe the image iconoscope is widely used. Of the latter there are British and French versions and also one that has been developed in the Philips Laboratory at Eindhoven (Netherlands). This Philips image iconoscope is described here and compared with other camera tubes.

The object of television is to transmit moving pictures via electrical means. This is achieved by "measuring" in succession the brightness of the very large number of picture elements into which the picture to be transmitted is imagined as being divided. This measuring consists in the conversion of the brightnesses into corresponding fluctuations of an electric current which in some way or other govern the signal transmitted.

To be reproduced with a satisfactory definition a picture has to be divided into some hundreds of thousands of elements; let us say, for the sake of convenience, that there are 400,000. Just as in cinema projection, to keep the picture free of flicker the number of pictures produced per second has to be greater than a certain minimum; in Europe the picture-repetition rate is mostly 25 complete pictures per second.

Hence the time available for measuring the brightness of one picture element is only $^1/_{400,000 \times 25}$ sec $= 10^{-7}$ sec. Already in 1908 — when 10^7 picture elements per second was beyond the wildest dreams and one had in mind a number something like 160,000 — Campbell Swinton realized that it would be imperative to have a practically inertialess electronic apparatus [1]).

To explain what is meant by this, let us describe the so-called flying-spot method frequently used for televising film pictures.

On the fluorescent screen of a cathode-ray tube (*fig. 1*) a raster of lines is scanned. A lens L_1 focuses this raster on the film picture to be transmitted, and another lens L_2 projects the transmitted light onto a photo-electric cell (a multiplier tube). The latter produces a current varying in strength according to the density of the successively illuminated picture elements on the film. The fluctuating current forms a signal with which the intensity of the electron beam in a cathode-ray tube (picture tube) at the receiving end is modulated. This beam is synchronized with that at the transmitting end. Thus the televised film picture is reproduced on the screen of the picture tube.

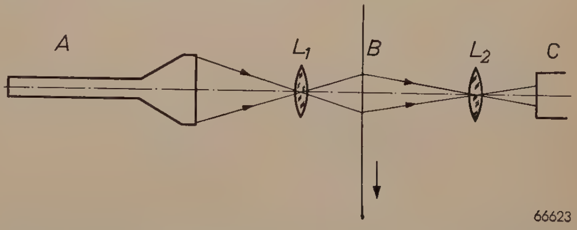


Fig. 1. Flying-spot method for televising film pictures. *A* cathode-ray tube with fluorescent screen on which a raster of lines is scanned. The lens L_1 projects this raster onto the film picture *B* and the transmitted light is concentrated by the lens L_2 in the multiplier tube *C*.

In principle it is possible to televise any scene in a similar way, by projecting a raster of light onto the scene and picking up the scattered light in a photocell. Practice shows, however, that in such a case the photo-electric currents are so small as to be "drowned" in the statistical noise of the photo-electrons, which noise creates at the receiving end the impression of the scene having been televised during a snowstorm! This poor result is due to the fact that only a very small part of the light scattered by the scene in all directions can be picked up in the photocell, such being contrary to the case with the film picture, where practically all the transmitted light can be collected in the photocell.

Fundamental failings of this otherwise fairly simple method are that the fluorescent spot is a weak source of light and that each picture element is illuminated for only 10^{-7} sec, so that one has to manage with the naturally limited amount of light thereby obtained.

---

[1]) A. A. Campbell Swinton, Distant electric vision, Nature **78**, 151, 18th June 1908, where a remarkable outline is given of the principles of present-day television.

However, another method can be imagined, whereby the scene is illuminated continuously, while for each picture element in succession in the space of time of 10^{-7} sec a signal is transmitted which corresponds to an illumination that was present during the whole of the time ($^1/_{25}$th sec) available per picture. This idea is to be found materialized in all present-day television camera tubes (with the exception of the tubes used for transmitting film pictures according to the method outlined above). With this method there is a continuous accumulation of charge during a frame period, and thus these tubes have come to be known as "storage tubes".

The oldest form of storage tube is the iconoscope, designed by Zworykin (1933), which will presently be dealt with. In the main this article will be devoted to a modern camera tube named the image iconoscope. Some other types will be mentioned in passing.

Classification of modern camera tubes

In the most important camera tubes of modern design there is a plate ("target" or "mosaic") on which is projected an electrical image, consisting of a two-dimensional pattern of electric potentials corresponding in amplitude and position to the luminance in the optical image of the scene to be transmitted. This electrical image is scanned point by point by a focused beam of electrons (the scanning beam), the potentials being thereby reduced to a certain "stabilizing potential" which in some way or other produces an electric signal.

The target is, of course, made of an insulating material, e.g. mica. When an electron beam is directed upon it the rule is that for every surface element, in the stable state, on an average just as many electrons have to be emitted as impinge upon it. This number of electrons may be zero or greater than zero, and it is these two possibilities which, as will be seen, form the basis for the classification of camera tubes into two groups.

Behaviour of an insulator when bombarded with electrons

Let us consider the following case. A beam of electrons (the primary electrons) is focused upon an insulating plate set up in vacuo. In the vicinity of the plate is an electrode, the collector, which has an adjustable positive potential V_{coll}.

When the primary electrons impinge upon a surface element of the plate they release secondary electrons from the material. The secondary-emission coefficient δ, i.e. the average number of secondary electrons released by one primary electron, depends upon the material and the velocity (thus the energy) of the primary electrons at the plate. If V_{coll} is so high that the collector attracts all the secondary electrons towards it then the variation of δ as a function of the energy V_{pr} (expressed in electron-volts) of the primary electrons is as represented in *fig. 2*. In the case of most materials there are two values for V_{pr} where $\delta = 1$; the smaller of the two is denoted by V_1, the larger by V_2. For mica, for instance, these material constants are in the order of 10 volts and some thousands of volts respectively.

As already stated, there are two possible stable states of the surface element, one where no electrons strike or leave the surface, and one where per second a certain number of electrons (>0) impinge upon it and just as many leave it. Which of these two states will be obtained depends upon the value of V_{pr}. If V_{pr} is less than V_1, and consequently δ is smaller than 1, the number of primary electrons striking the surface element per second will be greater than the number of secondary electrons released from it. Thus the potential of the element drops; only when it has reached zero (the cathode potential) — no primary electrons at all then reach the plate — is one of the two stable states of equilibrium obtained.

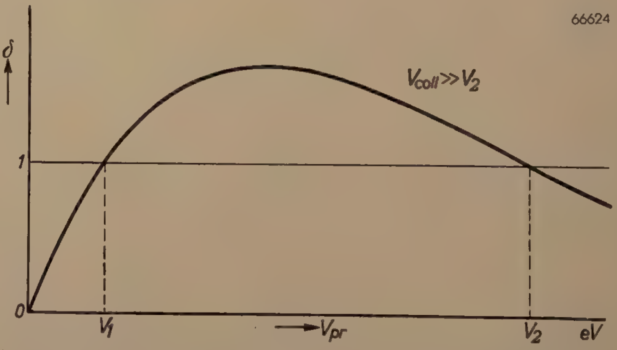


Fig. 2. Secondary-emission coefficient δ of an insulator, as a function of of the energy V_{pr} of the primary electrons. At two values of V_{pr} (V_1 and V_2) δ is equal to 1.

This curve applies when the collector potential is high compared with V_2.

Let us now consider the case where V_{pr} lies somewhere between V_1 and V_2. Then δ is greater than 1, the number of secondary electrons released from the surface per second is greater than the number of primary electrons impinging upon it, and the potential of the surface rises until at V_2 a stable state is reached.

The supposition that all the secondary electrons pass over to the collector only holds if V_{coll} is much higher than V_2. If that is the case then, upon gradually reducing V_{coll}, the shape of the curve does not change at first until V_{coll} approaches V_2. This is explained by the fact that when the surface element has a potential higher than that of the collector the released secondary electrons have to overcome an electric field before reaching the collector. Whether they succeed in doing so, or whether they fall back upon the target, depends entirely upon their energy and the geometry of the set-up. The potential of the surface will be stabilized at a value V_3, where the current intensity of the secondary electrons actually reaching the collector (i_{coll}) is equal to the current intensity i_{pr} of the primary beam. The ratio $i_{\text{coll}}/i_{\text{pr}}$ may be called the effective secondary-emission coefficient, δ_{eff}. For $V_{\text{pr}} < V_3$ (but $> V_1$) δ_{eff} is greater than unity, and for $V_{\text{pr}} > V_3$ the effective secondary-emission coefficient δ_{eff} is less than unity. As a rule V_3 is slightly higher than V_{coll} (*fig. 3*); in contrast with V_1 and V_2, V_3 is therefore not a material constant.

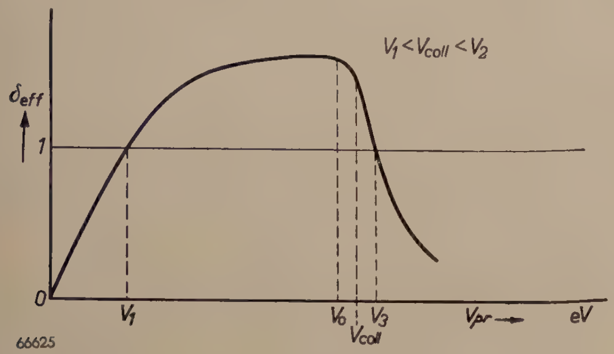


Fig. 3. Effective secondary-emission coefficient δ_{eff} of an insulator, as a function of the energy V_{pr} of the primary electrons when the collector potential V_{coll} is smaller than V_2 (cf fig. 2). $\delta_{\text{eff}} = 1$ at $V_{\text{pr}} = V_1$ and at $V_{\text{pr}} = V_3$, the latter value being a few volts higher than V_{coll}. To the left of V_0 (slightly lower than V_{coll}) the curve is identical to that in fig. 2.

Thus it is seen that when bombarded with slow electrons ($V_{\text{pr}} < V_1$) the surface potential becomes stabilized at zero, and when bombarded with electrons of high velocity it becomes stabilized at the value V_2 (provided $V_{\text{coll}} > V_2$) or at V_3 ($\approx V_{\text{coll}} < V_2$). For the target of a camera tube however no use is made of the value V_2, for practical reasons; it is strongly influenced by the condition of the surface and thus is too variable from point to point.

It is according to these possibilities that camera tubes are classified as:

1) low velocity tubes, where the target is stabilized at cathode potential (on that account they are also referred to as "cathode potential stabilized tubes", or for short "CPS tubes"), and

2) high velocity tubes, where the target is stabilized at the potential $V_3 \approx V_{\text{coll}}$ (e.g. 1000 V).

Among the first belongs the image orthicon, which is the type of tube mainly used in the U.S.A., while belonging to the second group are the iconoscope and the image iconoscope, the latter often being given preference in European television circles. One of the reasons for this preference is related to the large number of lines adopted on the West-European continent (625, and in France 819): with a high electron velocity it is easier to satisfy the high requirements for the the focusing of the scanning beam which are demanded for the definition required for such a large number of lines.

In this article exclusively high velocity tubes will be dealt with, though at the end some comparisons will be made with the image orthicon.

The iconoscope

The iconoscope is the camera tube which at the time gave such an impetus to television[2]). It is schematically represented in *fig. 4*, while in *fig. 5* a photograph is given of the Philips iconoscope, type 5852.

A lens (objective) projects an image of the scene onto a target of thin mica coated on the front with a mosaic of minute, mutually insulated, photo-sensitive elements. On the reverse side is a coating of metal, called the signal plate, forming the output electrode and externally connected to earth via a resistor. A ring-shaped coating of metal on the inside of the envelope serves as collector and is connected to earth direct.

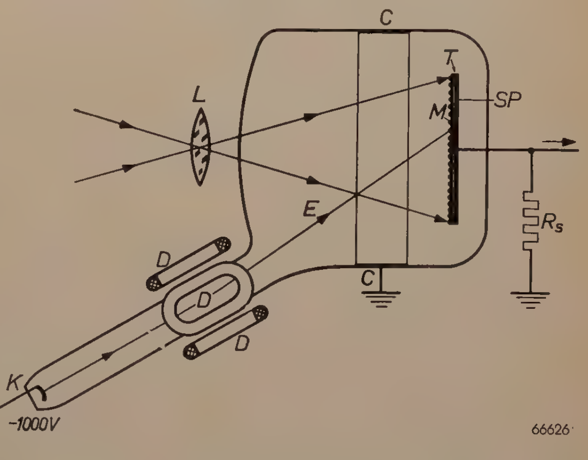


Fig. 4. Iconoscope. L a lens projecting the scene on the mosaic M of the target T. SP signal plate, R_s load resistor, C collector, K cathode, D deflection coils, E scanning beam. The (electrostatic) focusing is not shown.

In an arm of the envelope is an electron gun supplying a beam of electrons which is focused on the target plate. The cathode is at a potential of, say, —1000 V with respect to earth. The beam is made

---

[2]) See, e.g., Philips techn. Rev. 1, 18-19, 1936.

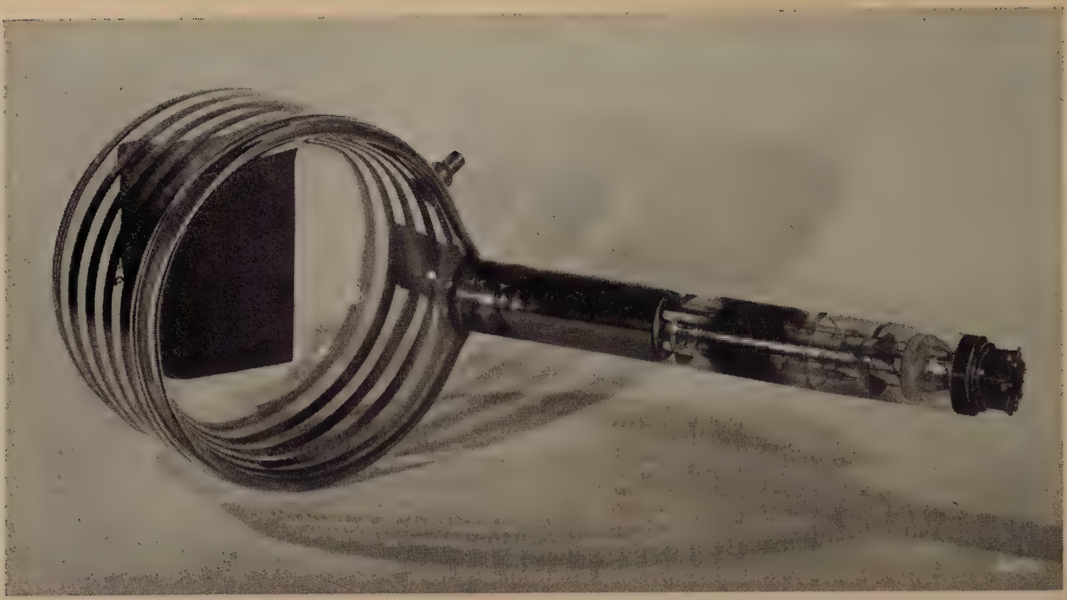


Fig. 5. The Philips iconoscope, type 5852.

to scan the mosaic continuously by means of two pairs of deflection coils. Thus a picture element is formed by the small area of the surface of the mosaic struck by the beam; although this is very small it covers a large number of photo-electric elements.

The action of the iconoscope is sometimes explained in the following (inadequate) way. The incident light causes the photo-electric elements of the mosaic to emit photo-electrons, which are taken up by the collector. Thus a positive electrical image is formed on the mosaic. The photo-electric elements together with the target form as many minute capacitors. As the scanning beam moves across the mosaic the group of capacitors belonging to a certain picture element are discharged. Through the resistor via which the signal plate is earthed there then flows a small current corresponding in intensity to the charge of the picture element, thus corresponding to the local luminance of the optical picture. Thus in the scanning of the electrical image a series of current impulses are generated which together form the video current.

Actually the position is not so simple as this. Such a description does not take into account the part played by secondary emission[3]). In point of fact we have here again the case of an insulator bombarded with electrons, where the collector potential is smaller than V_2 (cf. fig. 3). For the moment photo-emission may be left out of consideration by assuming the iconoscope to be in total darkness.

Not all the secondary electrons reach the collector, firstly because the potential of a bombarded surface element is higher than V_{coll}. If the scanning beam were to be brought to a standstill then, as explained with reference to fig. 3, the element would assume a potential V_3, for instance, 3 volts higher than V_{coll}. Normally, however, the beam is in motion and "rests" on the surface element for such a short space of time (10^{-7} sec) that although the value actually reached (V_{max}, *fig. 6*) is a fraction of a volt lower than V_3, it is still always higher than V_{coll} and thus prevents secondary electrons from reaching the collector. Furthermore, the collector is rather far removed from the target.

The secondary electrons which do not reach the collector fall back on other parts of the mosaic,

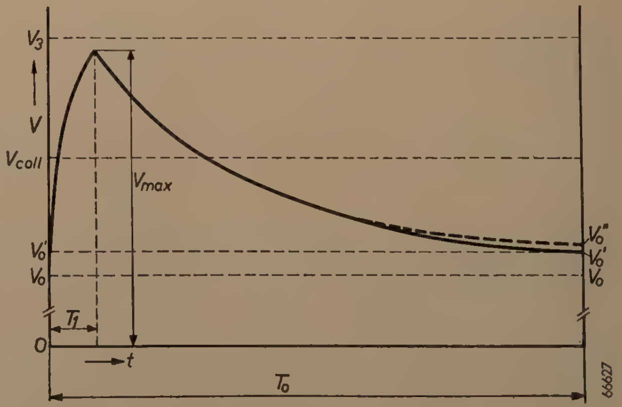


Fig. 6. Curve of the potential V of a picture element on the mosaic of an iconoscope, as a function of the time t. Fully drawn line: mosaic not illuminated; broken line: mosaic illuminated. T_0 = scanning period for the whole image ($^1/_{25}$th sec), T_1 = scanning time for one picture element (10^{-7} sec; in the drawing highly exaggerated). For the meaning of V_{coll} and V_3 see fig. 3, and for V_0, V_0', and V_0'' see the text.

---

[3]) V. K. Zworykin, G. A. Morton and L. F. Flory, Proc. Inst. Rad. Engrs 25, 1071-1092, 1937, and W. Heimann and K. Wemheuer, Z. techn. Phys. 19, 451-454, 1938.

which, as we have already seen, are at a slightly higher potential than the collector. This distribution of the secondary electrons is called the redistribution effect, and it is of essential importance for the action of the iconoscope.

After the surface element in question has been scanned, thereby assuming the potential V_{max}, it will continue to receive secondary electrons originating from other surface elements, until it is scanned by the beam again. Thus its potential V begins to drop (fig. 6) and, if a sufficient length of time were to elapse before the next scanning, would ultimately reach a value V_0 so low that the surface element in question could no longer take up secondary electrons from its surroundings. In practice, however, the surface element is again struck by the beam $^1/_{25}$th sec after the first scanning; the potential it reaches immediately prior to the second scanning is denoted by V_0' (fig. 6).

How is this course of affairs affected when the mosaic is illuminated? An illuminated image element will not emit photo-electrons continuously: from fig. 6 it is seen that during a considerable part of the scanning period the potential V of the element is higher than V_{coll}, and the photo-electrons do not possess sufficient energy to overcome this potential difference. Photo-emission begins, therefore, when — owing to the redistribution effect — the potential V has been sufficiently reduced. Experience teaches even that only during about the last 5% of the scanning period is V low enough for the photo current to reach the point of saturation. Owing to the photo-emission the right-hand part of the V curve of the picture element in question comes to lie somewhat higher (see the broken line in fig. 6), ending with a value $V_0'' > V_0'$, so that in the next scanning the surge from V_0'' to V_{max} is smaller than that from V_0' to V_{max} in the case of a non-illuminated element. It is this difference that gives rise to the signal current. (The illumination causes a change also in V_{max}, but this is small compared with the difference between V_0'' and V_0'.)

The most important features of the iconoscope will now be briefly discussed.

Sensitivity

As already explained, it is due to the redistribution effect that photo-emission can take place, but this is only possible during a fraction of a scanning period. Thus we are still far removed from a continuous photo-emission such as was imagined in the case of an ideal storage tube! This is one of the reasons for the rather low sensitivity of the iconoscope.

A second cause of the lack of sensitivity lies in the mosaic form of the light-sensitive layer. The insulation between the elements does not contribute towards photo-emission, so that a considerable part of the surface of the target is photo-electrically inactive.

Spurious signals

Another drawback attaching to the iconoscope is the fact that, in the case of the non-illuminated tube, the potential surge from V_0' to V_{max} when the element is scanned (fig. 6) varies with the position on the mosaic. Thus a certain signal, a "spurious" signal, is given also when the iconoscope is not illuminated.

The main cause of spurious signals (see the literature quoted in footnote [3]) is that the redistribution does not take place in the same way all over the mosaic, owing to the surroundings of the elements not being the same everywhere. When a certain line is being scanned the part of the mosaic above that line was more recently scanned than the part below it and thus has higher potentials. This is manifested in the picture at the receiving end showing a luminance increasing from top to bottom.

There are similar differences between the left and right sides and between the edge and the middle: the side of the picture where the lines end is reproduced with a higher luminance than the side where they begin, and the edge is lighter than the middle of the picture.

An additional effect is the variation, taking place during the scanning, in the angle at which the beam strikes the mosaic. This angular variation is accompanied by a variation both in the size of the area of the picture element and in the secondary emission coefficient.

When the iconoscope is illuminated the spurious signal is superposed on the picture signal and only if the latter is of a reasonable strength is the spurious signal not very disturbing. It is for this reason that with the iconoscope very high intensities of illumination are needed.

Linearity

The stronger the illumination on a certain part of the mosaic, the higher is the potential V_0'' at that spot just before it is scanned by the beam. This has two consequences: there is slightly less chance of further photo-electrons escaping, and there is a somewhat greater attraction of redistributed secondary electrons. Both these effects result in the amplitude of the signal increasing less than proportionately with the illumination. This nonlinearity is rather an advantage than a disadvantage in that it compensates fairly well an inverse nonlinearity between the beam current and the control voltage in the picture tube of the receiver. Thus there is no need to take steps to compensate the latter non-linear effect.

Other features

Apart from the non-linearity there are some other favourable features of the iconoscope, namely that the pictures obtained with it are good both in geometrical configuration and in gradation, and that its installation and operation do not involve any particular technical difficulties.

The image iconoscope

The greatest disadvantage of the iconoscope is its lack of sensitivity, and it is for that reason that attempts have been made to develop camera tubes with greater sensitivity, while still retaining the good picture quality obtained with the iconoscope when the scene is sufficiently illuminated.

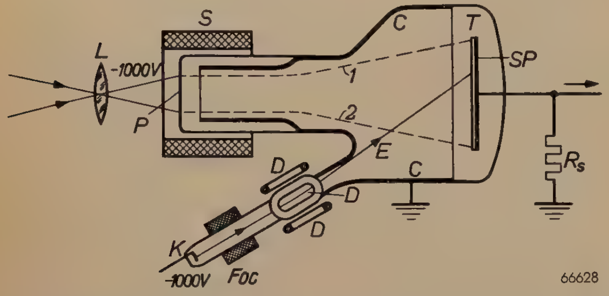


Fig. 7. Image iconoscope. *P* photocathode, *S* coil of the magnetic electron lens, *1* and *2* paths of photo-electrons, *Foc* focusing coil. Other letters have the same meaning as in fig. 4.

A year or two prior to 1940 a more sensitive version of the iconoscope, called the image iconoscope, was developed in the U.K. and in the U.S.A. [4]). Some improvements on this have since been made in the Philips Laboratory at Eindhoven, as will appear in the course of this article.

In the case of the image iconoscope (*fig.* 7) a lens (objective) projects an optical image of the scene to be televised onto a continuous, transparent photocathode. The local density of emission of the photo-electrons corresponds to the local luminance of the optical image. This photoemission image is focused by an electron lens onto a target consisting in this case of a thin layer of insulating material applied to the signal plate. The metallized inner wall of the envelope serves as collector. An electron gun mounted in an arm of the envelope supplies the beam of electrons scanning the target.

The differences, compared with the conventional iconoscope, which are mainly responsible for the gain in sensitivity, are the following:

1) The surface of the photocathode is continuous, so that none of its effective area is lost in insulation between the separate photo-electric elements.
2) The stream of photo-electrons reaching the target is reinforced by secondary emission, each photo-electron releasing on an average more than two secondary electrons.
3) The secondary electrons released from the target by the photo-electrons have a much greater energy than the photo-electrons in the ordinary iconoscope, so that secondary emission from a surface element begins immediately after that element has been stabilized by the scanning beam. This means a considerable gain in storage action. The curve for the potential V of a surface element is therefore similar to that in fig. 6, except that in the case where the cathode is illuminated the curve begins to diverge earlier from the curve applying for a non-illuminated cathode, the difference between V_0'' and V_0' thus being greater.

Let us now consider more closely the principal parts of the image iconoscope and also the important question of electron-optical projection. The Philips type of image iconoscope is illustrated in *fig. 8*.

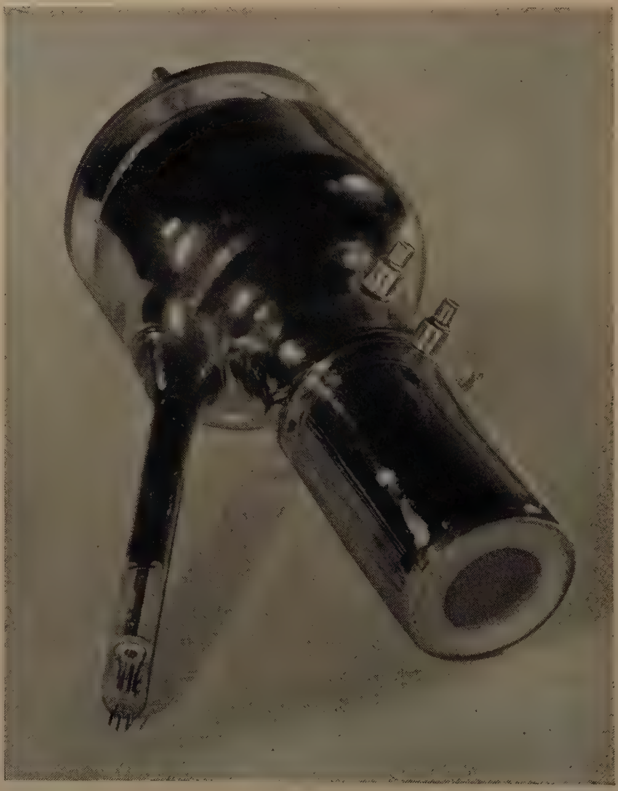


Fig. 8. The Philips image iconoscope, type 5854.

---

[4]) See, e.g., H. Iams, G. A. Morton and V. K. Zworykin, The image iconoscope, Proc. Inst. Rad. Engrs. **27**, 541-547, 1939.

The photocathode

Contrary to ordinary photo-electric cells, an image iconoscope must have a photocathode which is semi-transparent, because the light enters from the outside while the photo-electrons have to emerge on the inside.

In addition to this transparency great sensitivity is needed (at least some tens of μA photo-current per lumen incident light) and moreover the spectral sensitivity must not deviate too much from the relative luminosity curve of the average human eye.

These three requirements greatly restrict the choice of photo-electric material to be used. The photocathode in the Philips image iconoscope consists of a very thin coating of cesium, antimony and oxygen applied to a flat part of the glass envelope. The sensitivity for the light from an incandescent lamp with colour temperature 2600 °K is about 45 μA per lumen. The spectral sensitivity curve, compared with the relative luminosity curve for the normal eye, is slightly displaced towards the blue (*fig. 9*).

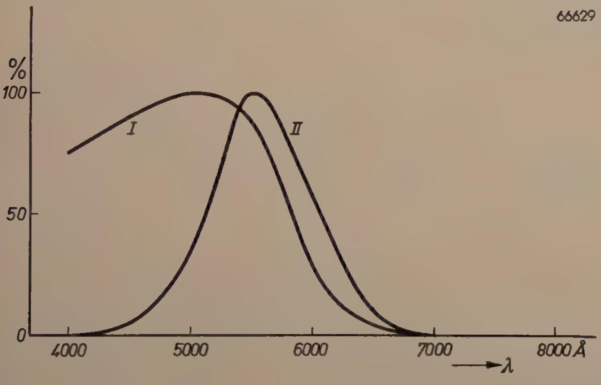


Fig. 9. Relative spectral sensitivity of the type 5854 image iconoscope (curve *I*), compared with the relative luminosity curve (*II*), as functions of the wavelength λ of the light.

Electron-optical image formation

The optical image of the scene is converted into a corresponding photo-emission image on the photocathode. The next step is to produce on the target an electrical image which is a faithful replica of the photo-emission image. This requires that the small beams of photo-electrons emitted from points of the photocathode are focused on corresponding points on the target. For this electron-optical image formation an electron lens is needed. As such use can be made of electric and/or magnetic fields.

For the image iconoscope the choice of the nature of the electron lens is determined mainly by the following considerations:

1) For the emission of photo-electrons to reach the saturation point an accelerating electric field is needed at the photocathode.

2) For a faithful image of a large part of the cathode surface to be produced in a not too long tube, the useful visual field angle of the electron lens — i.e. the angle from which the edges of the image are seen from the centre of the electron lens — must be large.

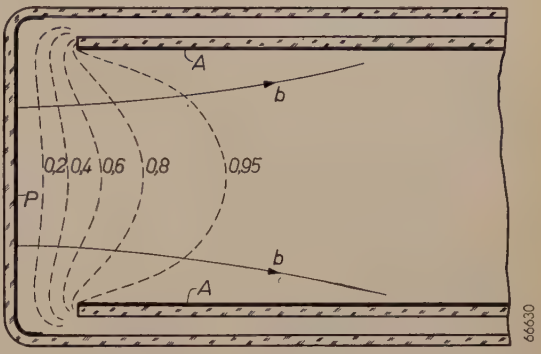


Fig. 10. Between the photocathode *P* and the metal coating *A* (on the inner wall of a glass cylinder) is an accelerating electric field. Broken lines: equipotential planes (the potential of *P* is zero, and that of *A* is taken as 1). *b* are electron paths, in the absence of other fields.

3) The image formed must be sufficiently sharp at all points. A prerequisite is that both the cathode and the target are flat (if the photocathode were curved it would be necessary to employ means of optical correction for the proper projection of a scene upon it; a curved target is very difficult to make and when scanned would cause barrel distortion).

Condition (1) implies that an electric field has to employed. This is obtained by means of a metal cylinder (e.g. the metal coating on the inner wall of a glass tube, *fig. 10*) facing the photocathode and applying a potential difference of, say, 1000 V between these electrodes. Since the cylinder forms, electrically, one whole with the earthed collector, the photocathode is given a potential of —1000 V with respect to earth.

This electric field alone, however, does not suffice. Fig. 10 shows some equipotential planes and some electron paths, from which it is seen that the latter diverge, so that we have here a negative electron lens. Now a negative lens cannot produce a real image from a real object. Therefore, in order to obtain a positive lens, either the shape of the electric field has to be changed in some particular way or a magnetic field has to be added which focuses each electron pencil.

The first alternative comes into conflict with the requirements (2) and (3), for it has not yet been found possible to design an electrostatic lens system which is capable of forming a faithful image of a sufficiently large part of a flat photocathode on a flat target.

The addition of a magnetic field is therefore indicated. Such a field can be produced by means of a coil placed concentrically around the tube. The coil has to be of such dimensions and in such a position as to minimize aberrations, whilst the magnetic field must not disturb the movement of the scanning beam. As the best solution a fairly long coil was used (see the article quoted in footnote [4])), protruding beyond the end of the tube (see *fig. 11*) far enough for the plane of the photocathode to be intersected practically at right angles by a uniform magnetic flux.

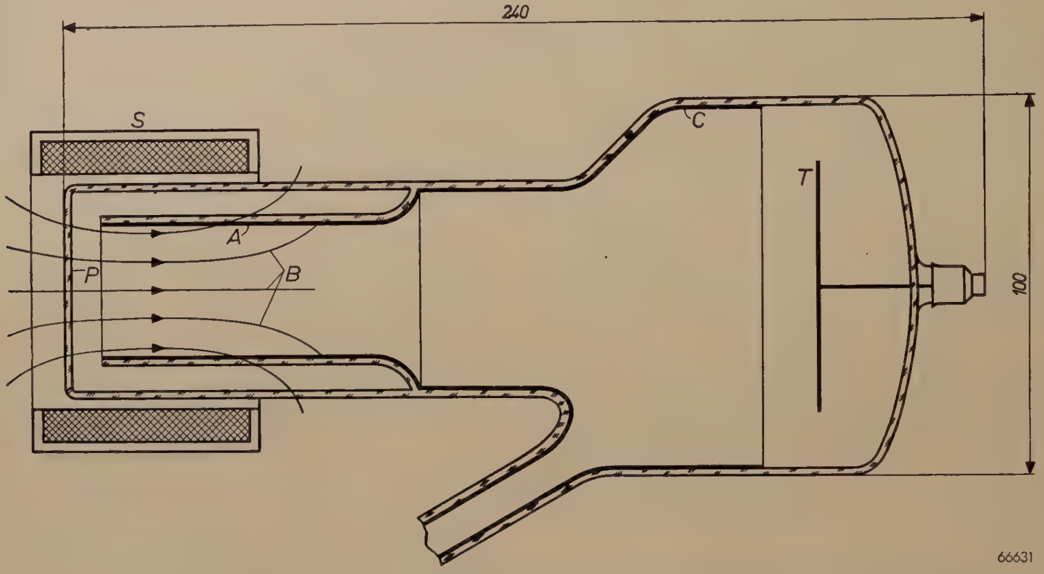


Fig. 11. Formation of the electron-optical image of the photocathode P on the target T with the aid of an electric field (between P and the cylinder A, cf. fig. 10) and a magnetic field. The latter (lines of flux density B) is produced by a focusing coil S.

For a theoretical treatment of the mechanism reference may be made to a publication elsewhere [5]), but the results of this study may be summarized and qualitatively commented upon here.

The movement of the electrons depends not only upon the two fields mentioned but also upon the velocities of the electrons leaving the photocathode. Some of them have zero initial velocity, and the paths they follow are called the principal rays. Generally, however, the electrons leave the cathode with a certain velocity, with the result that they follow a more complex path.

In order to get some idea of the principal rays we have to consider what electrostatic and magnetic forces act upon an electron leaving the cathode with zero velocity (e.g. at the point M in *fig. 12*.)

The cathode being an equipotential plane, the electric lines of force (field strength F) are perpendicular to it. Provided the cathode is at a sufficient depth inside the coil, the same may be said to be approximately the case with the lines of magnetic flux density B. Thus, under the influence of the electric field the electrons following a principal path emerge from the cathode at right angles to its plane; at first they are not subject to any force from the magnetic field.

Very soon, however, there will be a noticeable divergence of the F lines, and a force directed radially outward deflects the path of the electrons outward (fig. 10), thereby causing it to intersect B lines before these diverge. The electron is then subjected to a tangential Lorentzian force, which, with the direction of the B lines as drawn in fig. 11 and for an electron travelling above the centre line, is directed forward. Thus a rotation about the axis of the tube (the z axis) is superposed on the movement of the electron along the lines denoted by b in fig. 10; viewed in the direction of the positive z axis (fig. 12), the electron therefore rotates clockwise. Consequently the principal ray is no longer a planar curve but describes a sort of helix about the z axis. Owing to its having acquired tangential velocity the electron becomes subject to a secondary Lorentzian force directed radially towards the z axis and thus counteracting the divergence of the path.

---

[5]) J. C. Francken and R. Dorrestein, Paraxial image formation in the "magnetic" image iconoscope, Philips Res. Rep. 6, 323-346, 1951 (No. 5).

Towards the end of the coil the B lines diverge more with respect to the z axis than does the principal path of the electrons, so that there the electron becomes subject to a Lorentzian force which is tangential, like the first one, but directed backward instead of forward (for an electron above the z axis), because the angle between the B lines and the direction of motion is of opposite sign. Thus the clockwise rotation about the z axis is retarded and even reversed. Then a force comes into action which is directed radially away from the z axis.

Briefly, therefore, the course of a principal ray is as follows: at first the path is approximately parallel to the z axis, then it diverges farther and farther from that axis, thereby turning about the z axis first clockwise and later anti-clockwise in the form of a widening helix.

Although most of the electrons which leave the photocathode have velocities greater than zero and thus do not follow any principal paths, still it is the principal rays which determine the geometry of the electricon-optical image. Each forms the axis of a small electron pencil.

This may be explained as follows. Let us resolve the initial velocity into three components: a radial one (v_r) and a tangential one (v_φ) in the plane of the cathode, and an axial one (v_z) parallel to the z axis (see point N in fig. 12).

Electrons with initial velocity components $v_z = 0$ but v_r and $v_\varphi \neq 0$ follow a path more or less helical about the principal ray but ending, owing to the secondary Lorentzian forces, in the same point on the target as the principal ray. Thus the components v_r and v_φ do not cause any aberrations.

Such is not the case, however, with v_z. This component gives rise to a certain "chromatic" aberration: a point of the photocathode from which electrons emerge with axial velocity does not result in a point being formed on the target but a small circle (scattering circle), the diameter of which is:

$$d = \text{const.}\,\frac{\varepsilon}{F}\,,$$

where ε is defined by $e\varepsilon = \frac{1}{2}m\,v_0^2$ (e = charge, m = mass and v_0 = initial velocity of the electron), and F = electric field strength at the cathode. In the image iconoscope ε is in the order of 1 volt, F about $6 \cdot 10^4$ V/m. Owing to this high value of F the diameter d is so small — thus the image so sharp — that the image iconoscope can quite well be worked with more than 600 scanning lines. In the image orthicon, on the other hand, F is ten times smaller[6]), so that with this type of tube the formation of the electron-optical image is a limiting factor for the number of lines.

---

[6]) H. B. De Vore, Proc. Inst. Rad. Engrs **36**, 335-345, 1948.

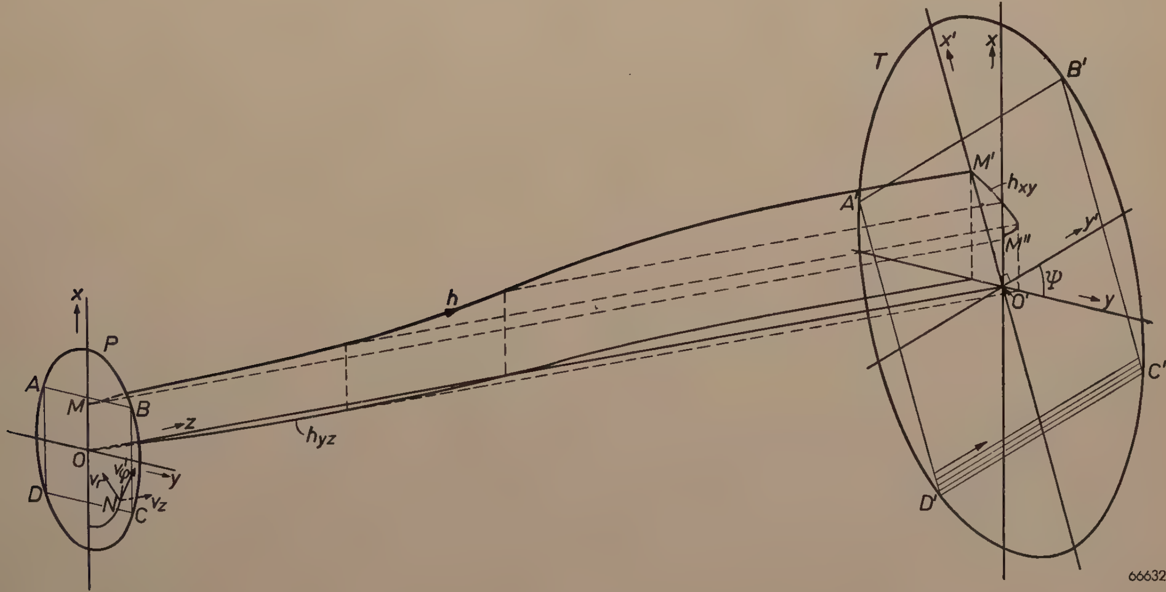


Fig. 12. Formation of the electron-optical image of the photocathode P on the target T. h is a principal path followed by a photo-electron emerging at M with zero initial velocity and reaching the target at M'. M'' is the projection of M on T, h_{xy} the projection of h on T, and h_{yz} the projection of h on the yz plane. At the point N on the photocathode are drawn the three components v_r, v_φ and v_z of the velocity of an electron emitted with a certain initial velocity.

The electrical image on the target is turned over an angle Ψ with respect to the image on the photocathode. $A'B'C'D'$ is the part of the target scanned (a few lines are drawn on it), $ABCD$ the corresponding part of the photocathode. The diagonal $A'C'$ is 75 mm long, while the length of the diagonal AC is normally 20 mm; when other coils are used AC is maximum 27 mm and minimum 10 mm.

Owing to the predominance of the diverging forces acting upon the electrons following the principal path the image on the target is magnified, and owing to the tangential forces the electron image is rotated with respect to the optical image on the photocathode, the angle of rotation Ψ (fig. 12) being about 30 to 40 degrees.

The degree of magnification is related to the effective length of the magnetic field: if the field is lengthened (in the direction of the target) it takes longer for a photo-electron to reach the area where the B lines appreciably diverge, and the secondary Lorentzian force radially directed towards the z axis is active longer, so that this converging force then has greater influence and hence the magnification is smaller. Conversely, the shorter the effective magnetic field, the greater is the magnification.

With our image iconoscope the magnification is normally 3.75, which means to say that the scanned part of the target, which always covers an area of 45 mm × 60 mm, corresponds to an area of 12 mm × 16 mm on the photocathode (the diameter of the active part of the photocathode is 20 mm). By exchanging the coil for another of different dimensions it is also possible, however, to work with a larger or a smaller magnification, thus projecting a smaller or a larger part of the photocathode on the target. The choice as regards the size of the effective photocathode is governed by requirements of an optical, light-technical and camera-technical nature. The limits for the magnification are 2.75 and 7.5 (diameter of the projected part of the cathode respectively 27 mm and 10 mm).

With a magnification greater than 7 to 8, owing to the "chromatic" aberration of the photo-electrons emerging with axial velocity (see above) there is too great a loss in resolving power.

The lower limit of 2.75 is due to various other aberrations, which with a smaller magnification can no longer be sufficiently compensated. As such may be distinguished: field curvature, pin-cushion distortion and so-called S distortion. The first two are known from light-optics [7]). By S distortion is meant the effect of the image of a straight line being projected as a line curved somewhat in the shape of the letter S *(fig. 13)*. This effect, which always occurs to a certain extent with magnetic lenses, is due to the fact that the principal rays at the edge of the image are rotated over a larger angle than those nearer the middle. As is the case with pin-cushion distortion, S distortion is proportional to the third power of the diameter of the active part of the photocathode, thus in the case where the size of the image on the target is constant this distortion is inversely proportional to the third power of the magnification. If the magnification is not too small the S distortion can be sufficiently corrected by electrical means (which we cannot enter into here), but if it is less than 2.75 this is no longer possible. In *fig. 14a* a picture is given showing all three aberrations to a marked extent. The picture in fig. 14*b*, however, has only a scarcely perceptible S distortion, which is not troublesome.

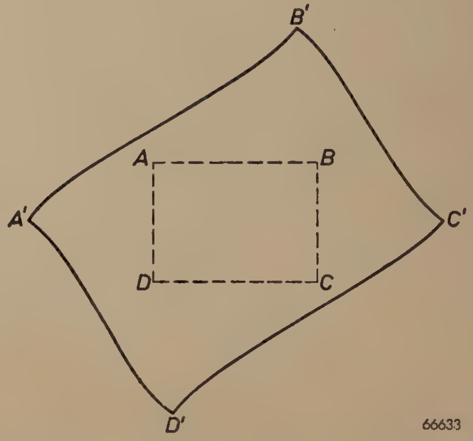


Fig. 13. *ABCD* is an image on the photocathode, *A'B'C'D'* the corresponding electrical image on the target. The latter is magnified and turned with respect to *ABCD* and also shows some S distortion, which always occurs when magnetic lenses are used (straight lines are projected with a slightly S-shaped curve). If the magnification is too small the S distortion becomes so pronounced that it can no longer be sufficiently corrected.

The effective length of the magnetic field can, in principle, be varied in a simpler way than by exchanging the coil, namely by shifting the coil along the tube. In practice, however, the magnification can be only slightly varied in this way: as the coil is shifted farther towards the target a strong field curvature very soon arises (the B lines are then no longer at right angles to the photocathode), whilst if it is slightly shifted in the other direction it touches the objective (to give the camera a large visual field an objective with a short focal length is used, and it is therefore close to the photocathode).

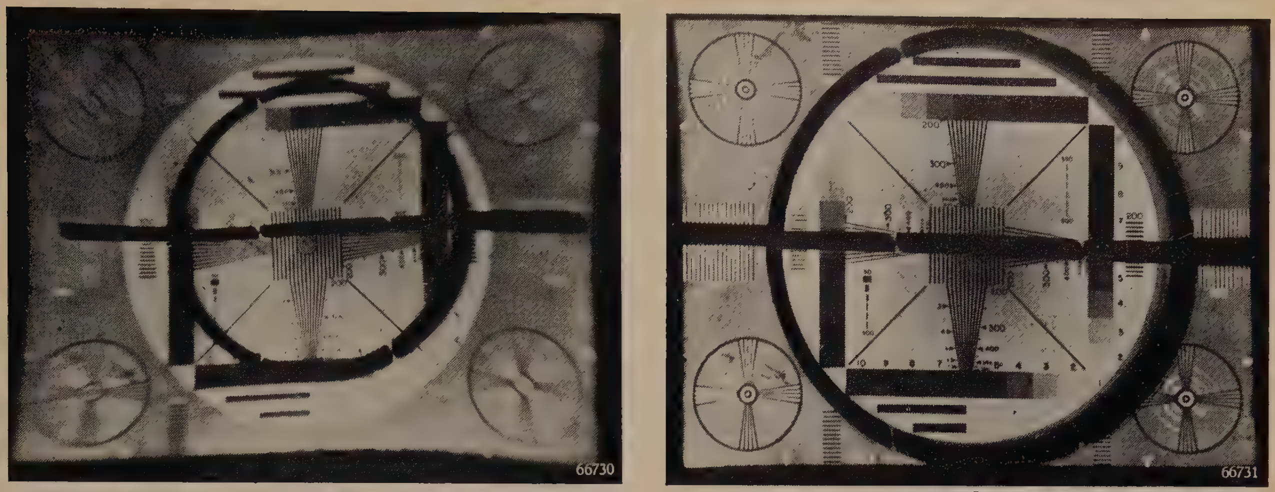
The electron gun

The electron gun supplies the scanning beam. Just as is the case with most picture tubes, in the image iconoscope the beam is focused and deflected with the aid of magnetic fields.

In regard to the sharpness of the scanning, there are two things to be considered. First let us deal with the non-deflected beam. This is focused on the centre of the target, where its diameter must be so small that the lines do not overlap when being scanned. If it is desired to work for instance with 1000 lines then, if the height of the scanned part of the target is 45 mm, the effective diameter of the focus must not be more than 45 μ. This requirement is all the better fulfilled the higher the acceleration voltage is chosen, but, as will be shown in the next section, this voltage should preferably not exceed 1000 V.

---

[7]) A review of various optical aberrations is to be found, for instance, in: W. de Groot, Philips techn. Rev. 9, 301-308, 1947, in particular pages 304 and 306.



a *b*

Fig. 14. *a*) Picture showing a marked field curvature, pin-cushion distortion and S distortion. *b*) Here there is only a slight S distortion, which can easily be corrected electrically.
These photographs have been taken with the aid of an experimental tube in which a fluorescent screen was used instead of a target. On the photocathode a test pattern was projected as used in television for detecting aberrations and checking the definition and gradation. The heavy black circle and the thick horizontal line in the middle correspond to markings on the photocathode for determining the magnification.

Further, account has to be taken of the fact that in the image iconoscope the electron gun has to be mounted with its axis at an angle to the target. Consequently when the beam is deflected upward or downward the focus is no longer situated on the target. Therefore, to obtain sufficiently sharp scanning also away from the centre, the beam must have a good depth of focus, which means that is has to be extremely narrow. Hence the angle of divergence $2\alpha_f$ (see *fig. 15*) has to be kept very small.

It is, in general, difficult to obtain a fine focus with a very narrow beam on account of the mutual repulsion of the electrons, but fortunately the intensity of the beam current required is very low, in the order of 0.1 μA.

In addition to this space-charge repulsion there is another factor limiting the spot size obtained with a very narrow beam: there is a very fundamental relationship between the angle of divergence $2\alpha_f$ and the current density in the beam. In the case where the space charge is negligible this relationship is:

$$\sin^2 \alpha_f = \frac{V_0}{V} \cdot \frac{j_f}{j_0}, \quad . \; . \; . \; . \; . \quad (1)$$

where $V_0 = {}^1/_2\, m v_0{}^2/e$ (with m = mass, v_0 = initial velocity and e = charge of an electron), V = the potential difference traversed by the electrons, j_f = density of the beam current in the focus, and j_0 that at the cathode of the gun.

Equation (1) is derived as follows. In optics there is a law (ascribed to Abbe as well as to Helmholtz or to Lagrange) which says that the brightness of a picture divided by the square of the index of refraction of the space in which the picture is formed is constant (as long as losses through absorption and reflection are negligible). In electron-optics the analogue of this law is expressed in the following formula:

$$\frac{j}{V \sin^2 \alpha} = \text{const.}, \quad . \; . \; . \; . \; . \; . \; . \; . \; . \quad (2)$$

where j is the current density and α half the angle of divergence of the electron beam, while V is the potential difference between the point in question and the (imaginary) point where the electron velocity is zero.

The values of these quantities at the cathode are denoted by j_0, α_0 and V_0. The value of α_0 is 90°, because the electrons emerge from the cathode in all directions. (The same holds if there is a limiting diaphragm in the beam, as is the case in the gun of our image iconoscope, since the electron spot is formed by small beams of electrons originating from different

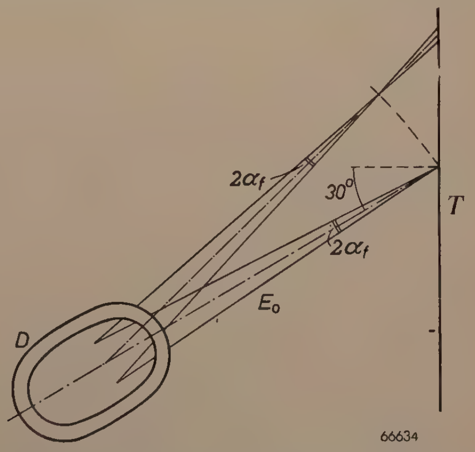


Fig. 15. Assuming that the non-deflected beam E_0 has been focused onto the centre of the target T, when the beam is deflected the focus will no longer be in the plane of T. This gives rise to blurring, the extent of which increases with the angle of divergence $2\alpha_f$.

points of the cathode and each having a half-angle of divergence of 90° at the cathode [8]). Thus the value of the constant in (2) is j_0/V_0.

For the electrons which in the gun of the image iconoscope have traversed a potential difference V the relationship between the half-angle of divergence α_f and the current density j_f in the focus is given by (2) as:

$$\sin^2 \alpha_f = \frac{V_0}{V} \cdot \frac{j_f}{j_0},$$

which is the same as equation (1).

From eq. (1) it is seen that, with given values of V_0, V and j_0, a reduction of α_f leads to a reduction of j_f and thus, for a given value of the beam current intensity, to a larger cross section of the focus. Further it appears that the greater the value of j_0 (thus the heavier the cathode may be loaded) and the higher the value of V, the less is the danger of the focus becoming too large.

What has to be found, therefore, is an optimum value for α_f at which, on the one hand, the focus is not too large and, on the other hand, the sharpness at the edges of the image does not differ too much from that in the middle. With our image iconoscope the position is such that this optimum value of α_f lies at about 3×10^{-3} radians.

This small angle of divergence, combined with a low beam current intensity (about 0.2 μA), has been obtained by placing two diaphragms in the beam. The first, with a narrow aperture, confines the beam within the desired small angle. The second one, with a wider aperture, allows the beam to pass through without hindrance but intercepts the low-velocity secondary electrons formed round the edge of the first diaphragm.

With the focus of 45 μ already mentioned (1000 lines at 45 mm) and a beam current of 0.2 μA, the average current density in the focus is $j_f = 12$ mA/cm^2. Substituting this in eq. (1), and for V the value at which the secondary emission coefficient of the target is greatest, viz. 1000 V, and for V_0 the value corresponding to the average initial velocity ($\approx$ 0.1 V), we find for the average current density at the cathode of the gun $j_0 \approx 120$ mA/cm^2. The peak value of the current density is in fact several times greater. Although an ordinary oxide-coated cathode may indeed be continuously loaded with such a current density, it is better to use what is known as an L cathode [9]), since this has a much longer life. It would be quite undesirable if the useful life of a costly tube such as the image iconoscope were to be dependant upon the life of a component like the cathode of the gun.

From the manufacturing point of view the L cathode also has the advantage that it lends itself well for mounting with very close tolerances of the distance between cathode and grid. Thanks to this property there are fewer rejects, and thus less wastage, in manufacture through errors of mounting in the gun.

The glass arm of the envelope containing the electron gun has been kept as narrow as possible (internal diameter 11 mm, external 14 mm), so that also the focusing coil and the deflection coils may be small. The advantages of this are : (1) that the field of the deflection coils in the space facing the target is so weak as not to have any perceptible influence upon the paths followed by the photo-electrons, and (2) little power is needed for excitation of the coils, so that the generators of the deflection currents, mounted together with the tube in the camera, can also be small. This makes the handling of the camera all the easier. A camera with an image iconoscope is shown in *fig. 16*.

In practical use the resolving power of the Philips image iconoscope is found to be 900 to 1000 lines in the middle of the image and about 700 lines at the edges. (These limits are set by the electron gun; the resolving power of the electron-optical projection is very much greater.)

The target

As in the conventional iconoscope, also in the first of our image iconoscopes the target was of mica (of course without mosaic). In the preparation of the photocathode cesium vapour is driven into the tube. This vapour tends to affect the surface of the mica, thereby causing the secondary emission coefficient, δ, as a function of the energy V_{pr} of the primary electrons impinging on it, to follow a curve as represented by *I* in *fig. 17*. It is true that the peak of this curve lies higher than in the case where the mica has not been in contact with cesium, but then this peak is situated at a point corresponding to a fairly small energy of the impinging electrons. For the electron-optical projection the electric field at the photocathode should be strong, and this, as already stated, is reached with a potential difference of 1000 V. In the electron gun, as we have seen, a high accelerating voltage is necessary in order to obtain a fine focus; for the sake of simplicity a potential difference of 1000 V is likewise chosen here. With an energy of 1000 eV, however, δ remains considerably below the maximum value.

---

[8]) D. B. Langmuir, Proc. Inst. Rad. Engrs 25, 982, 1937.

[9]) H. J. Lemmens, M. J. Jansen and R. Loosjes, A new thermionic cathode for heavy loads, Philips techn. Rev. 11, 341-350, 1950.

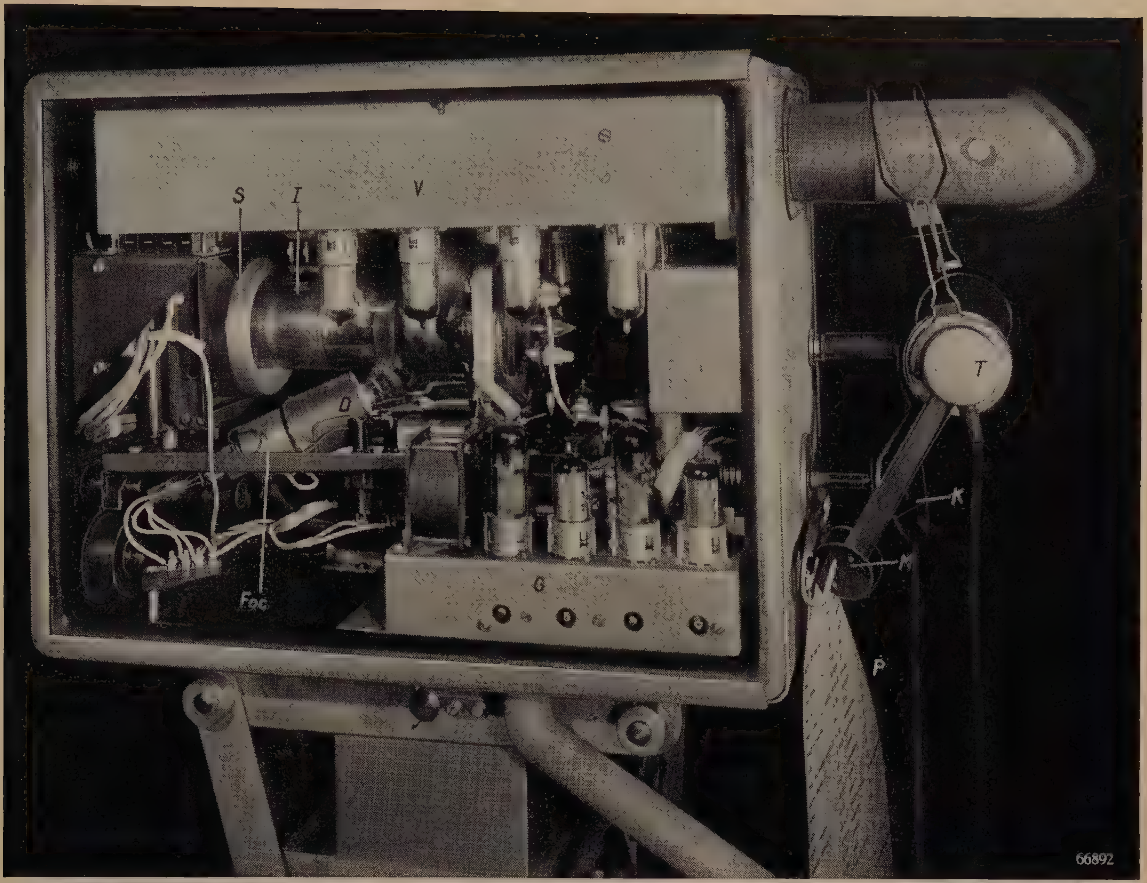


Fig. 16. One of the cameras used for the experimental television broadcasts at Eindhoven. One side panel and a screen have been removed. *I* image iconoscope, type 5854. *S* image coil, *Foc* focusing coil. *D* deflection coils. *G* time-base generator. *V* chassis with monitor picture tube and accessories. *M* microphone and *T* telephone for communication between the operator and the control room. *K* knob for exchanging the objective. *P* playbook.

An improvement has been reached by coating the mica with a thin layer of MgO. The curve $\delta = f(V_{pr})$ now has the shape of curve *II* in fig. 17: at $V_{pr} \approx 1000$ V the coefficient δ is about 4, which means a considerable gain in sensitivity. Furthermore, owing to the coating of MgO, stains on the mica which cannot be removed and otherwise show up clearly in the picture are thereby made invisible.

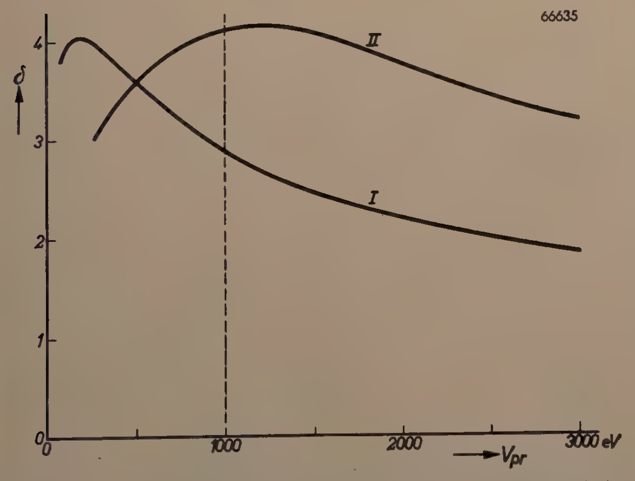


Fig. 17. Secondary-emission coefficient, δ, as a function of the energy V_{pr} of the primary electrons, *I* for mica affected by Cs, *II* for mica coated with MgO. At $V_{pr} \approx 1000$ V *II* is much more favourable than *I*.

Notwithstanding this gain, the sensitivity of the first tubes made was still unsatisfactory. A suitable method of measuring the sensitivity of the target — i.e. the ratio of the signal current to the electron current striking the target — is the following. Except for a narrow vertical strip in the middle, which is kept dark, the photocathode is uniformly illuminated with a variable intensity of light. Each time the scanning beam passes one of the edges of the "black" strip on the target the signal current suddenly changes. The magnitude of this current surge I_s at the transition from "black" to "illuminated" is measured as a function of the photocurrent I_{ph}, which is varied with the strength of illumination.

The result, at first, was that the signal did not rise sufficiently above the noise level of the first amplifying valve. In order to gain an insight into the factors determining the shape of the signal curve an attempt was made to approach theoretically the problem of the stabilization of the potentials on the target and the production of the signal. Suffice it to say here that these investigations [10]) led to the conclusion that the capacitance of a surface element of the target with respect to the signal plate is an important factor, and that an increase of this capacitance must lead to greater sensitivity.

New tubes were therefore made with a mica sheet of only about 25 μ thickness (also with a layer of MgO, thin compared with the mica) instead of the original sheet thickness of 50 μ.

The success attained from these two measures — applying a layer of MgO and halving the thickness of the mica — is evident when comparing curve *II* in *fig. 18* with curve *I*.

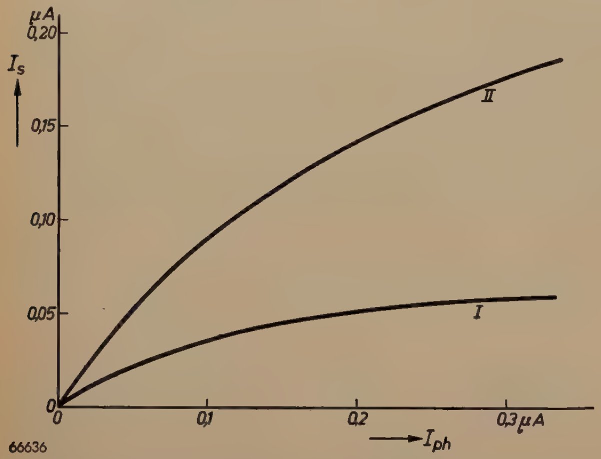


Fig. 18. The current surge I_s in the signal current each time the scanning beam passes from a "black" strip on the target to an "illuminated" part, plotted as a function of the photo-current I_{ph}. Curve *I* applies for a mica sheet 50μ thick without MgO, curve *II* for a sheet 25μ thick coated with MgO.

The reproduced picture of a scene televised under the normal studio lighting, or of an outdoor scene in daylight (even in bad weather), with the image iconoscope last described, is almost free of "noise" and shows excellent gradation.

The scanning beam has to erase, as it were, the electric image on the target point by point. For it to be able to do so sufficiently also on a thinner target, the strength of the beam current had to be increased approximately in proportion to the capacitance of the target. The value of 0.2 μA mentioned corresponds to the mica thickness of 25 μ.

In the image iconoscope spurious signals arise from the same cause as in the case of the conventional iconoscope: the various surface elements of the target are not all in the same position with respect to the scanning beam. In the image iconoscope, however, the situation is more favourable: with the tube described (mica 25 μ thick, beam current 0.2 μA) and with an illumination producing a photo-current of more than 0.1 μA, the spurious signals are so weak that there is hardly any need of compensating measures. In practice a photo-current of 0.1 μA can be obtained with an illumination of the scene of about 1000 lux, when using a non-diaphragmed, normal objective with aperture *f*:2.

Comparison of different types of camera tubes

Let us now compare, briefly, the two main types of camera tubes, the high-velocity and the low-velocity types.

In the first place there is the question of sensitivity. This resolves itself into two factors (disregarding the efficiency of the optical system), viz. the sensitivity of the photocathode (photo-current I_{ph} in relation to the light flux falling on the cathode) and the sensitivity of the scanning mechanism (ratio of signal current I_s to photo-current I_{ph}).

As regards the sensitivity of the photocathode it has already been seen that of the two described representatives of high-velocity tubes — the conventional iconoscope and the image iconoscope — the latter has very much the advantage, owing to the continuity of the photocathode. Among the low-velocity tubes there are likewise types with a mosaic cathode and others with a continuous cathode, the latter including the image orthicon, which as regards photocathode sensitivity is equal to the image iconoscope.

The scanning sensitivity of low-velocity tubes (and without an inportant improvement to be mentioned below) can be directly determined: in these tubes the electrons emitted by the light-sensitive layer are supplemented by electrons from the scanning beam, the scanning sensitivity in this case therefore being exactly 1 μA signal current per μA photo-current. In high-velocity tubes the phenomenon of redistribution complicates matters, but from measurements taken the scanning sensitivity of the ordinary iconoscope can be put at $^1/_{20}$th μA/μA and that of the image iconoscope, with weak to moderate illumination, at about 1 μA/μA (see fig. 18).

Although, therefore, the image iconoscope has about the same scanning sensitivity as the (simple) low-velocity tube, there is a difference in the $I_s = f(I_{ph})$ curve in favour of the image iconoscope: in the case of this tube, and also the ordinary iconoscope, the said curve is not linear (see fig. 18), whereas in the case of low-velocity tubes it is linear; the non-linear curve is favourable, as explained when dealing with the iconoscope.

---

[10]) P. Schagen, On the mechanism of high-velocity target stabilization and the mode of operation of television-camera tubes of the image-iconoscope type, Philips Res. Rep. **6**, 135-152, 1951 (No. 2).

There is a means, however, of appreciably increasing the scanning sensitivity of low-velocity tubes. The electrons from the scanning beam which are not taken up by the target and return to the gun can be collected in a multiplier, placed around the gun, which works with secondary emission and thus multiplies them. This is what takes place in the image orthicon, the camera tube commonly employed in the U.S.A. In this way the scanning sensitivity may be raised to a value of 25 to 100 μA/μA, which is of course a valuable property when, for instance, scenes have to be televised in poor light (e.g. sporting events under artificial light). An inherent drawback is, however, that the current of the returning beam can be modulated only up to about 20% and consequently contains a relatively large amount of noise. More electrons return from dark parts than from light ones, so that the noise from the darker parts is both relatively and absolutely the strongest, and often very troublesome.

It has already been explained that in regard to spurious signals the image iconoscope has a decided advantage over the ordinary iconoscope. The image orthicon is free of spurious signals of this nature, but on the other hand it is subject to another interference connected with the fact that the secondary-emitting surfaces of the multiplier do not have exactly the same secondary-emission coefficient over the whole area ("dynode spots").

Electron-optically, high-velocity tubes have undeniably the advantage over those of the other group, in that with electrons of a high velocity it is easier to obtain a scanning beam with a high resolving power, and there is much less trouble from interfering electric and magnetic fields.

Summary. In the introduction television camera tubes are divided into two groups, high-velocity and low-velocity tubes. This article deals only with tubes of the former group, first briefly touching upon the iconoscope and then considering the image iconoscope in more detail.

Discussed in succession are the photocathode, the mechanism for the electron-optical projection, the electron gun and the target as found in the Philips image iconoscope, type 5854. Measured under the illumination from an incandescent lamp with colour temperature 2600 °K, the photocathode (Cs-Sb-O) gives an output of about 45 μA per lumen. For the formation of the electron-optical image an electric and a magnetic field are used, situated in the same space (near the photocathode) and giving a magnified replica of the photocathode image on the target. The electron gun supplies a magnetically focused and magnetically deflected scanning beam (beam-current strength normally 0.2 μA) with such a fine focus and such a focus depth that the resolving power at the middle of the target amounts to 900-1000 lines and at the edges is still 700 lines. The cathode employed in the gun is of the "L" type, which has a long life even under a high specific load. As regards the target it has been found that increasing the capacitance with respect to the signal plate gives greater sensitivity. A very thin mica sheet (25μ) has therefore been used, coated with a thin layer of MgO in order to increase the number of secondary electrons released by the high-velocity electrons impinging on the target ($\delta > 4$).

The article concludes with a brief comparison between high-velocity tubes and low-velocity tubes, and in particular between the new image iconoscope described and the image orthicon.

NEW DEVELOPMENTS IN THE GRAMOPHONE WORLD

by L. ALONS.

681.844.081.48

Developments of the last few years have been of great importance for gramophone enthusiasts, in that it has been made possible to play the whole of a classical work without interruption or at most with only one interruption at a suitable place in the composition. Now this greater enjoyment can be had from music at no additional cost. Drastic changes have, of course, had to be made both in the records and in the record players.

Introduction

As is known, gramophone records as are normally obtainable in the trade have to answer certain standards as regards dimensions, speed of rotation, etc. These standards were fixed at a time when almost exclusively mechanical sound boxes were used, but since these have now for the greater part been replaced by electrical pick-ups it has become possible to revise those old standards. An incentive in this direction was given by the appearance of the magnetophone, a magnetic-tape recorder and player with a long playing time, which threatened to become a competitor of the gramophone. This led to the development of a new kind of gramophone record — the long-playing record — which, in the 12″ size, has a playing time of $22^1/_2$ minutes on each side; compared with the $4^1/_2$ minutes of a normal record of the same diameter this is indeed a great advance, especially when it is considered that, in addition, the quality of the reproduction is vastly improved in many respects.

For dance music and for short musical pieces in general a long playing time is of less importance, but also for this category some producers have taken advantage of the possibility to depart from the old standards. The records now being made for such short pieces of music have a diameter of only 7″ for a playing time of $4^1/_2$ minutes. A striking feature about these "easy-storage" records, as they have come to be called, and also the long-playing ones is the large number of extremely fine grooves cut in them, and that is why both these new kinds of records are often given the name of microgroove records.

In this article only the long-playing records will be discussed, but similar considerations apply also for the other types of microgroove records. First the factors which set a limit to the playing time of a gramophone record will be considered, from which it will be seen that one of the conditions for extending the playing time was that the pick-up has to be made as light as possible. This article will conclude with a description of the new record players developed by Philips. These are fitted with an extremely light pick-up and reproduce with a high degree of fidelity the sound recorded both on long-playing records and on those of the old standards, whilst at the same time record wear is reduced to a minimum.

Factors limiting the playing time

The gramophone records obtainable in the trade are made by moulding a thermoplastic material in a matrix which is an electrolytically produced negative of an original recording on wax or lacquer. The original is obtained in the following way.

A metal disc covered with a layer of wax or lacquer is rotated at a constant speed under a cutter travelling at a uniform speed along a straight line from the circumference to the centre of the disc. In this way a spiral groove with a certain pitch is cut in the disc (the unmodulated groove). Now when the sound that is to be recorded is converted into radial oscillations of the cutter and these oscillations are superimposed upon the uniform movement of the cutter towards the centre, the modulations are cut in the disc and thus a modulated groove is obtained.

Obviously the playing time T depends upon the number of turns N of the groove cut and the number of revolutions per second R made by the disc. Hence:

$$T = \frac{N}{R}.$$

The number of turns can be expressed by equating the diameter of the outermost turn $D_{\max}$, the diameter of the innermost turn $D_{\min}$ and the pitch in the formula:

$$N = \frac{D_{\max} - D_{\min}}{2 \cdot \text{pitch}}.$$

and thus for the playing time we arrive at the expression:

$$T = \frac{D_{\mathrm{max}} - D_{\mathrm{min}}}{2 \cdot \mathrm{pitch}} \frac{1}{R} \quad . \; . \; . \; . \; . \; . \; . \quad (1)$$

If it is desired to increase the playing time without increasing the outer diameter of the record, in principle there are the following possibilities:

(1) Reducing the pitch.

Since the pitch is the sum of the width of the groove itself and the width of the wall between the grooves (pitch = wall width + groove width) the pitch can be reduced in two ways:
a) by reducing the width of the groove, or
b) by reducing the width of the wall.

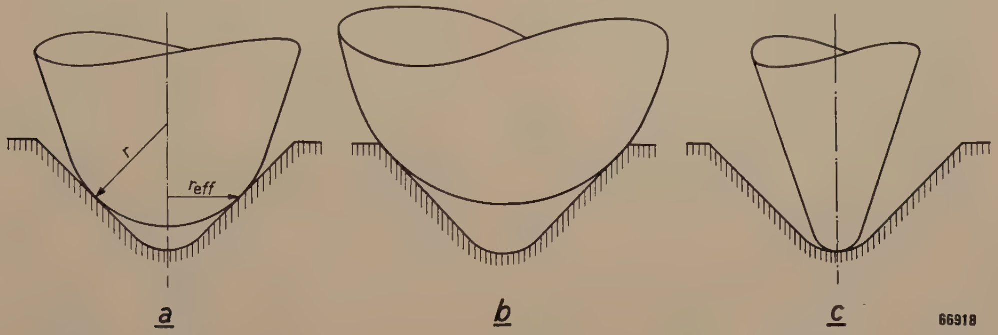


Fig. 1. Position of the needle in the groove. (*a*) How the spherically rounded tip of the needle should rest against the walls of the groove; *r* is the radius of curvature of the needle tip, r_{eff} the effective needle radius. (*b*) and (*c*) How it should not be. In the situation (*b*) the needle rests on the sides of the groove, which are subject to damage both in the cutting of the groove and in the further handling of the record; consequently with the needle in such a position there would be excessive noise. In the situation (*c*) the needle rides on the bed of the groove, and the forces required for "controlling" the needle can be produced only after it has run up against a side wall; this gives rise to play of the needle and thus distortion, in addition to excessive noise.

(2) and (3) Reducing either the speed of rotation or D_{min}. Anticipating the discussion of these possibilities, it is to be pointed out here that a relationship exists between R, D_{min} and the minimum groove velocity V_{m}, i.e. the linear velocity at which the innermost groove passes under the cutter or the playing needle, this being expressed as:

$$V_{\mathrm{m}} = \pi D_{\mathrm{min}} R \quad . \; . \; . \; . \; . \; . \quad (2)$$

It will appear that V_{m}, and not D_{min} or R, is the factor which counts.

Each of these possible measures will now be considered in turn.

The groove width

The lower limit of the groove width depends upon the properties of the material from which the ultimate gramophone record is made, as will be understood from what follows.

A narrower groove requires a sharper needle (the point of the needle has to have a smaller radius of curvature). The dimensions of the needle must be such that the rounded point of the needle come to rest against the sides of the groove, so as to avoid the possibility of the point gliding along the upper edges of the groove, which, owing to the roughness of the edges, would cause excessive noise (see *fig. 1*). Incidentally it should be added that the point of the needle must not be allowed to glide along the bottom of the groove, because it would then be unable to follow the oscillations of the groove without some play, which would cause both distortion of the sound and excessive noise.

The sharper needle required for a narrower groove implies — given equal weight of the pick-up — higher needle pressure on the disc. Naturally the needle pressure must be kept below the critical pressure at which non-elastic deformation of the disc material takes place.

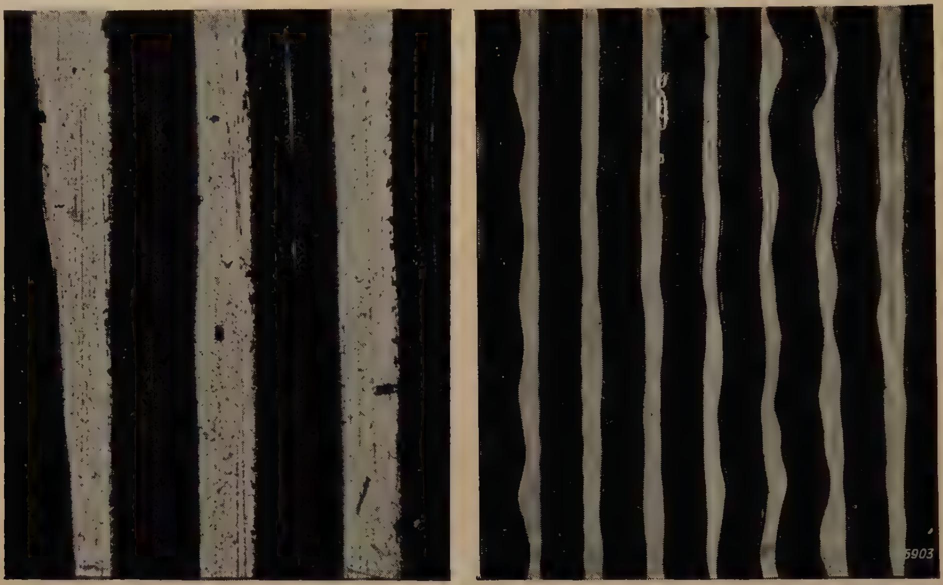
The groove width of long-playing records has been reduced, but this was possible only on account of the weight of the pick-up having been considerably reduced.

The width of the wall

The width of the wall is determined by the maximum amplitude, B_{max}, of the modulation made by the cutter. To avoid two adjacent modulated grooves from running into each other this width must be at least $2B_{\mathrm{max}}$. Consequently the width of the wall can be reduced only by reducing B_{max}, and the question is what factors set a limit to this.

A very rough reasoning — which, however, immediately brings us to the root of the matter — is this:

The material from which the gramophone record is made is not perfectly homogeneous but has a certain granular structure, so that there are irregularities in the groove, and these cause background noise. The sounds that are still to be heard clearly during reproduction must have an amplitude greater than the irregularities of the groove — thus a minimum "signal-to-noise ratio" is required. The amplitude $B_{\min}$ corresponding to the weakest sounds to be recorded is therefore directly dependent upon the degree of homogeneity of the disc material. The amplitude with which the sounds with the highest intensity are recorded, $B_{\max}$, must be sufficiently greater for proper expression to be given to the differences actually occurring between forte and piano passages (dynamic range). Bearing in mind that the thought from which we started was to make $B_{\max}$ as small as possible so as to obtain a longer playing time, it is evident that a compromise has to be made so as to satisfy both the artistic requirements and those for a not too short playing time.



a *b*

Fig. 2. Two pieces of gramophone records: *a*) with normal groove, *b*) with microgroove (enlarged 80 times). Note the smoothness of the surface in the second case.

Once an artistically acceptable relation has been found between $B_{\max}$ and $B_{\min}$, it is clear that the width of the wall can be reduced — and thus the playing time extended — provided a finer grained material is available for the records (see *fig. 2*).

Actually the matter is more complicated than this, since account has to be taken of the differences in frequency of the sounds to be recorded (from, for instance, $\omega/2\pi = 50$ c/s up to 10,000 c/s). There is no hard and fast rule for the manner in which sounds of equal intensity but differing in frequency are recorded on the disc — this is called the cutting characteristic — and great problems are still involved in the choice of a cutting characteristic satisfying all requirements. Considered in the light of what has been set forth above, the obvious thing to do would be to give these sounds one and the same amplitude. When reproduced with so-called amplitude pick-ups (where the voltage induced depends entirely upon the deflection of the needle) the sounds so recorded induce equal voltages which are converted in the loudspeaker into sounds of equal intensity. The consequence of such a cutting method "with constant amplitude" — assuming, for the sake of simplicity, that in the sounds to be recorded all frequencies have the same maximum intensity — would be that all frequencies on the disc would have the same recorded amplitude $B_{\max}$. This means that the accelerations to be given to the cutter for the highest frequencies ($\sim B_{\max}\,\omega^2$) would be so great as to be unattainable. (Such great accelerations would give rise to difficulties also in the playing back of the records, though these do not weigh quite so heavily as the difficulties encountered in the cutting, which are of a fundamental nature.)

Another obvious solution would be to adopt a cutting method for recording with "constant velocity": sounds of equal intensity are then recorded with the same amplitude of velocity $B\omega$, thus with an amplitude inversely proportional to ω. In the case of reproduction with a so-called velocity pick-up (equal velocity amplitude produces equal voltage) a faithful rendering of the sound recorded would then be obtained. But with such a cutting characteristic we are farther off than ever from attaining our object of reducing the width of the wall; because if for the highest frequency B_{min} is chosen as large as is necessary for an adequate signal-to-noise ratio and B_{max} is given a value such that the forces of acceleration then occurring are still permissible, the amplitudes for the low notes would assume impossible proportions. Moreover, for the highest notes the ratio $B_{\text{max}}:B_{\text{min}}$ would anyhow be smaller than that required for a good dynamic range.

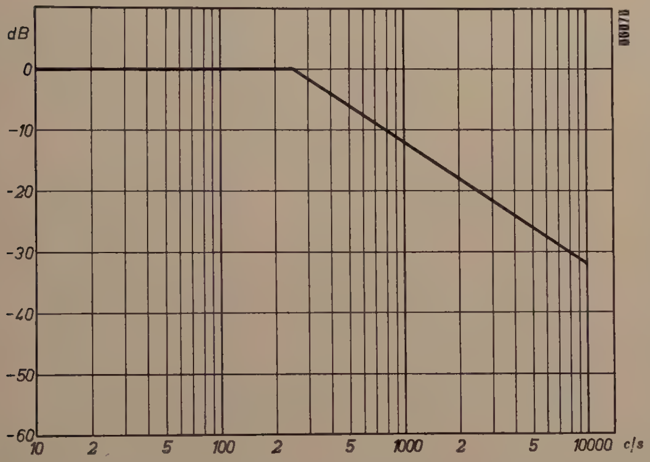


Fig. 3. Cutting characteristic of normal gramophone records. The line represents the amplitude of the modulation for tones of equal strength but with different frequencies, the amplitude being expressed in decibels. Apart from an additive constant the ordinates are therefore equal to 20 log B. Note the logarithmic scale on the abscissa axis. The cross-over frequency, where the knee occurs in the curve, differs slightly for records of different make.

In the cutting characteristics used in practice there is, therefore, always a compromise between the two methods mentioned; see *fig. 3*, where a cutting characteristic as hitherto commonly used is given by way of illustration. Though only amplitude pick-ups or velocity pick-ups may be available, this compromise does not in principle constitute any objection if the reproduction is by electrical means, since in that case the (linear) distortion which would occur in the reproduction can be avoided by using a suitable filter. Bearing in mind that a mechanical player with a gramophone horn is in principle a velocity pick-up, and that mechanical filters are not so easy to make, it can be understood why in the past the method of cutting with a constant velocity was adhered to as closely as possible.

This more precise approach to the problem of the maximum amplitude B_{max} that is to be chosen does not, however, detract from the truth of the conclusion derived from the rough reasoning first given. B_{min} is determined by the granulation of the disc material and B_{max} by the requirement that the ratio $B_{\text{max}}:B_{\text{min}}$ must not be too small, for the sake of reasonable dynamic range. The precise reasoning however shows that, after B_{max} has been fixed and thus also the playing time, the difference between forte and piano will not be given such good expression in the high frequencies as in the low ones.

The speed and the diameter of the innermost turn

The speed R and the diameter of the innermost turn D_{min} are limited by the fact that the minimum groove velocity V_{m}, to which they are both proportional (eq. 2), must not be below a certain limit, as otherwise the sound will be badly distorted and the record will suffer from excessive wear. This is related to two effects which are known as "tracing distortion" and "pinch effect". These two effects are in turn a direct consequence of the fact that the cutting and the playing back of the gramophone record are done with two instruments (the cutter and the needle) which have a different profile. The cutter must have a flat profile with cutting edges (see *fig. 4*), the groove being cut in such a way that the walls of a sinusoidal groove enclose a smaller angle (α) in the sides than in the peaks of the sine curve. The needle used for playing the record must

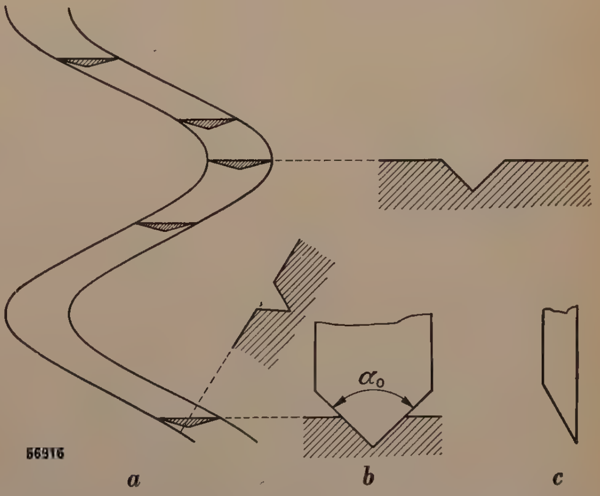


Fig. 4. *a*) Groove in which the cutter is represented in cross section at different places. The groove becomes narrower the more it deviates from the direction of the unmodulated groove. *b*) Front view of the cutter. *c*) Side view of the cutter. The angle α_0 is the angle given by the cutter to an unmodulated groove and to a modulated groove in the peaks of the sine. In this illustration and in fig. 6 the ratio of the modulation amplitude to the width of the groove is exaggerated (see fig. 2).

not, of course, have the same profile as that of the cutter, because otherwise each time the record is played it would, as it were, cut the groove again, which would result in prohibitive wear of the record. The needle should preferably have a spherically rounded tip and rest in the groove in the manner already described (see fig. 1).

It will now be shown what factors influence these two distortion effects.

Tracing distortion

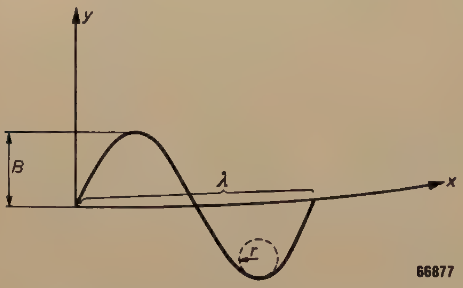


Fig. 5. The elements of the modulated groove. λ wavelength of the modulation, B amplitude of the modulation, r minimum radius of curvature of the modulation. The x-axis is the unmodulated groove.

Let us consider a short length of the unmodulated groove as being straight, with its centre line as x-axis (*fig. 5*). Denoting the direction perpendicular to the unmodulated groove in the plane of the disc as the y-direction, then for the centre line of a groove modulated with a pure sinusoidal tone we have the equation:

$$y = B \sin \frac{\omega x}{V} \quad \ldots\ldots \quad (3)$$

where ω is the angular frequency of the vibration recorded and V the groove speed of the turn in question. (For the innermost turn the groove speed is therefore $V = V_{\min}$.) The sine curve given by (3) is most strongly curved in the peaks, the radius of curvature there being approximately

$$r = \frac{V^2}{B\omega^2} \quad \ldots\ldots\ldots \quad (4)$$

Thus it is seen that the minimum radius of curvature decreases with decreasing groove speed. Now, as *fig. 6* shows, if the radius of curvature of the groove is comparable to the "effective" radius r_{eff} of the needle (i.e. the radius of that part of the needle point which rests against the walls of the groove; see fig.1) then the needle can no longer follow faithfully the modulation of the groove. The resultant distortion is called the tracing distortion. If this distortion is to be kept within a certain limit it necessarily implies that a limit is set for the groove speed of the innermost turn, viz. roughly $V_m^2 > B_{\max}\omega^2 r_{\text{eff}}$. If $B_{\max}\omega^2$, i.e. the acceleration at the highest frequencies, is chosen as large as is permissible on the foregoing grounds — this has been done for the long-playing records, as we shall presently see — then V_m can be reduced only if at the same time r_{eff} is also reduced, or, in other words, if a sharper needle is used.

Mention has already been made in this article of the favourable effect that a sharp needle has on the playing time, on account of this being related to a small groove width. It has also been pointed out that a sharper needle makes it essential to have a lighter pick-up so as to avoid non-elastic deformation of the record material resulting from too great a static needle pressure. There is, however, yet another factor, already mentioned in passing, namely the dynamic needle pressure, which results from the radial accelerations of the needle in the groove. The maximum value of these acclerations is $B_{\max}\omega^2$. It has already been stated that for the long-playing records this quantity has been chosen of the same order as that for normal records, and when it is pointed out that the dynamic pressures occurring in the playing of normal records with a normal pick-up certainly lie round about the limit of elasticity of the disc material, it is obvious that a limit is set for the sharpness of the needle. A sharper needle can be chosen only if the mass of

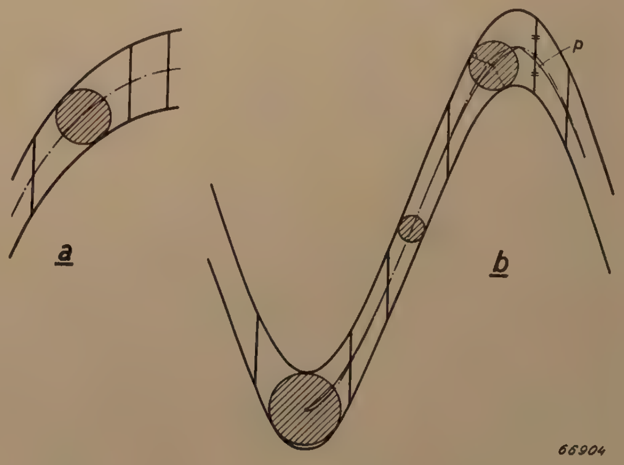


Fig. 6. How tracing distortion arises. *a*) The centre of the needle follows the centre line of the groove if the radius of curvature of the needle point is smaller than that of the groove. In *b*) the two radii of curvature in the peaks of the sine are equal. There the needle can no longer exactly follow the groove, as may be seen in the part of the curve indicated by *p*. The narrowing of the groove (reduced groove angle) along the side of the sine curve gives rise to pinch effect (see fig. 4). The cutter is represented schematically at different points.

the pick-up, in so far as it is rigidly coupled with the needle in respect to its horizontal movements, is made sufficiently small [1]).

Pinch effect

As already stated, the needle used for playing a gramophone record should bear against the sides of the groove. The height at which the point of the needle "rides" above the bottom of the groove depends upon the aforementioned groove angle α. Simultaneously with the periodical variations of α the needle together with the mass rigidly coupled to it is forced slightly upward (see figs 4 and 6). This periodical vertical movement of the needle is called the pinch effect. The variation in the height of the needle, and therefore the amplitude of the pinch effect, increases in direct proportion to the variation of the groove angle α. Now the steeper the slope of the sine curve the smaller is the groove angle α in the sides. As follows directly from equation (3), the steepest slope of the groove is given by:

$$\left(\frac{\mathrm{d}y}{\mathrm{d}x}\right)_0 = \frac{\omega B}{V} . \quad . \quad . \quad . \quad . \quad . \quad . \quad . \qquad (5)$$

Therefore the periodical variation of the groove angle α and consequently the amplitude of the pinch effect will increase with decreasing groove speed.

The frequency at which the pinch effect takes place — the "pinch frequency" — is twice the modulation frequency, since each sine curve has two peaks and two sides per cycle.

This pinch effect has two adverse consequences: (1) It leads to wear of the record, owing to the forces exerted by the needle on the record by the periodical vertical movement of the needle. It has already been shown that the periodical variation of the groove angle (to which the pinch effect is proportional) increases with decreasing minimum groove speed. The reduction of the latter quantity, which is what is aimed at, will therefore result in an increase of the pinch effect. Furthermore this effect will be greater still on account of the fact that a smaller V_{m} makes it necessary to use a finer needle (smaller contact surface); given equal force, this results in increased dynamic needle pressure. The only possible way to avoid greater wear of records having a smaller V_{m} is to keep as small as possible the mass m_1 which, in respect to the vertical movements, is rigidly coupled with the needle point. This is the course that has been successfully followed in the new Philips' pick-up which will be described later.

(2) The second consequence of the pinch effect is a certain distortion of the reproduced sound, due to the fact that in general the vertical movements of the needle also induce voltages in the pick-up. It will be seen that with the solution of the problem sub (1) this distortion can also be limited to a negligible minimum.

Long-playing records

It has already been said in what respects long-playing records differ from the old type, and we shall now go into these points in more detail.

Long-playing records are made from a new material named "Vinylite", which belongs to the family of plastics. It combines the advantages of a much smaller size of grain (see fig. 2) with unbreakability. It has already been explained how a smaller grain structure allows of a smaller width of wall and thus a smaller pitch of the groove.

In fact the groove pitch has been reduced from 0.26 mm in the old standard records to 0.1 mm, whilst the groove width has been narrowed down from 0.13 mm to 0.07 mm. At the same time the speed of rotation has been reduced from 78 r.p.m. to $33^1/_3$ r.p.m., and the diameter of the innermost turn has been increased from 4″ to 4.75″. Owing to these changes a playing time of $22^1/_2$ minutes has been reached. Compared with the old standard records the minimum groove velocity has been reduced by a factor of $\frac{33^1/_3}{78} \times \frac{12}{10} = 0.51$.

It might at first sight seem strange that the inner diameter should have been increased. It can be proved, however, that with a given minimum groove velocity there is an optimum value of the inner diameter and of the speed of rotation at which the playing time is the maximum. In other words, if a smaller inner diameter were chosen then, to keep the minimum groove velocity above the given limit, the speed of rotation would have to be increased to such an extent that in the ultimate the playing time would be shorter.

The playing time mentioned applies for the largest microgroove records having an outer diameter of 12″. In addition to these there are also smaller records of 10″ and the so-called "easy-storage" records, of only 7″ diameter and a playing time of $4^1/_2$ minutes. Not all microgroove records are played at a speed of $33^1/_3$ r.p.m., some records of 7″ having been made for a speed of 45 r.p.m.

---

[1]) When speaking here of certain masses being rigidly coupled to the needle in respect to movements in certain directions we mean certain effective masses which may be imagined as being rigidly connected to the needle. This occurrence of different effective masses for different directions of movement is related to the fact that various parts of the pick-up are elastically coupled to each other.

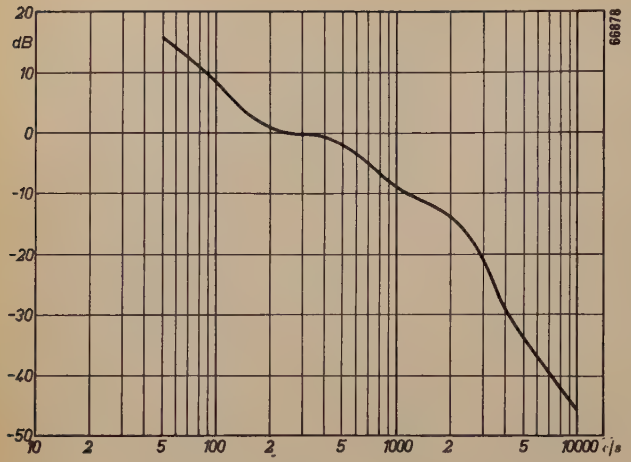


Fig. 7. Cutting characteristic of a long-playing record.

The cutting characteristic of a long-playing record is given in *fig. 7*, from which it is seen that over the whole range the amplitude increases with decreasing frequency. Thus the compromise between cutting at a constant velocity and cutting at a constant amplitude has been reached in a different way from that indicated in fig. 3.

It should be pointed out that, by choosing for $B_{\max}$ for the high tones a value of the same order as that of normal records, any possible gain in playing time that might be reached by further reducing V_m has been sacrified for the sake of better quality in the reproduction of the high frequencies.

Description of the pick-up used in the new Philips record player

All kinds of gramophone records, both the normal and the microgroove ones, can be played on the Philips type 2508 record player. This has one motor for three speeds, 78, 45 and $33^1/_3$ r.p.m., and further a record changer for automatically changing 12", 10" and 7" records.

A simpler design is the type 2978 record player (see *fig. 8*), which is without a record changer, while there is a choice of two speeds, 78 and $33^1/_3$ r.p.m.; the lack of the third speed however is not of any great consequence for all countries outside the U.S.A. since there are relatively few 45 r.p.m. microgroove records in these countries.

Both these types of record players are equipped with a new design of pick-up fitted with a specially made sapphire needle.

The needle

It appears that even to the trained ear a needle rounded off at the tip with a radius of curvature of 15 μ gives no perceptible tracing distortion, even on the innermost groove.

With a groove angle of 90° such a radius of curvature for the rounding off of the tip gives an effective needle radius (see fig. 1) of:

$$15 \cos \frac{a}{2} = 15 \cos 45^\circ = 11\ \mu.$$

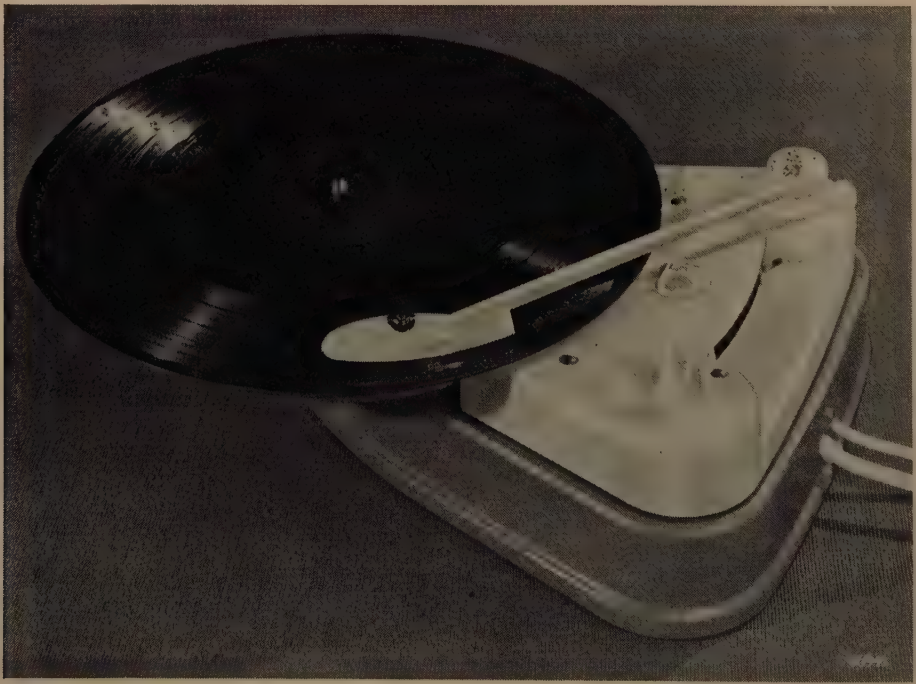


Fig. 8. The Philips type 2978 record player.

Such a fine needle point, 1 mm in length, can indeed be made, but then it could be used only for playing microgroove records, since on a normal record this fine point would ride on the bottom of the groove (see fig. 1*c*), which is not permissible. To be suitable for playing both microgroove and normal records the pick-up would therefore have to be fitted with two different needles, and as a matter of fact this is indeed done upon request for connoisseurs who want the reproduction to answer the very highest demands. For the normal type, however, it is desired to avoid this complication because if the gramophone were to be used thoughtlessly it might lead to trouble. A compromise has been found by making a needle which can be used for records with a normal groove as well as for those with a microgroove. The needle point is made of sapphire and has the shape of a truncated cone with the edges rounded off (see *fig 9*). The effective radius is 26μ, the total length 1mm, the thickness 0.4 mm and the weight 0.2 mg, i.e. approximately one-eighth of the weight of the sapphire needles hitherto used. Thus a needle has been produced which by its simplicity in use amply compensates the small loss in quality of reproduction of the high notes due to tracing distortion.

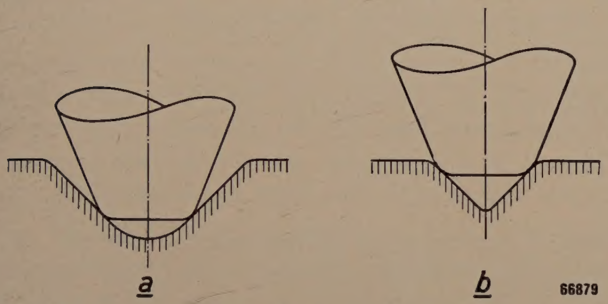


Fig. 9. The truncated sapphire needle in a normal groove (*a*) and in a microgroove (*b*). It is seen that the needle has been given such a shape and dimensions that while it does not ride on the bed of a normal groove it is nevertheless small enough to fit in a microgroove.

The piezo-electric crystal and the armature

The pick-up developed for playing microgroove records is of the piezo-electric type. This appeared to be the most suitable for a simple, light construction which will generate sufficiently high voltages without any additional amplification being required over and above that available in the A.F. part of a radio receiver.

This choice was made in spite of a certain prejudice felt against piezo-electric instruments on account of their susceptibility to atmospheric influences (e. g. very humid or extremely dry climates). Special measures have been taken against this in the packing of the pick-ups for the trade and care has also been taken to provide for adequate protection of the instrument while in use; these measures will be described later.

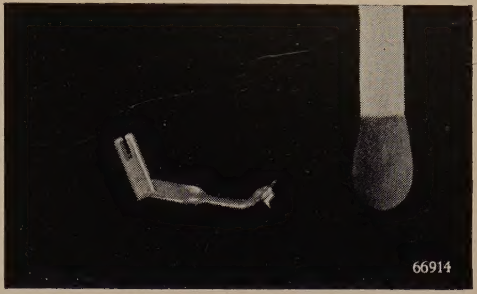


Fig. 10. The drag spring with needle. On the right the head of a match to give an idea of the dimensions.

The sapphire needle point is fixed to a drag spring (see *fig. 10*) connected to the piezo-electric crystal via the so-called armature (see *fig. 11*). Between the fork of the armature and the crystal is a piece of elastic, rubber-like material. The armature is rotatable about its longitudinal axis, with rubber rings serving as bearings. In this way horizontal deflections of the needle are translated into torsional movements of the crystal, which is so sawn out of a larger crystal as to react mainly to those movements. The part of the pick-up in which the crystal is contained is filled with a gel. The whole pick-up weighs no more than 15 grammes. The force exercised by the needle on the record is 7 grammes.

It has been pointed out already that not only the static needle pressure but also the dynamic pressures have to be kept sufficiently low. This means that the mass m_1, which plays a part in the pinch effect, has to be small because the frequency at which m_1 vibrates is twice the modulation frequency. Provision has therefore been made for this by connecting the needle to the armature via the drag spring, thereby coupling the needle to the crystal so loosely, as far as the vertical movements are concerned, that the vibrations of the needle due to the pinch effect are not conveyed to the crystal. The effective mass m_1 has thereby been reduced to a few milligrammes, and as a consequence no trouble is experienced from the above-mentioned distortion accompanying these vertical movements.

Owing to the accelerations taking place in the horizontal direction, the mass m rigidly connected with the needle in this direction has also to be kept small. This means primarily that the armature, which as a matter of fact also yields a small contribution towards m_1, has to be made as light as possible. How drastically the mass of the armature has been reduced may be judged from the fact that the new armature weighs only 30 mg, as compared with the 2500 mg of the armature formerly used, which makes an enormous difference (see *fig. 12*).

The function of the elastic layer between armature and crystal may be explained as follows: The rigidity of this material depends greatly upon the frequency at which it is deformed, increasing with increasing frequency, so that vibrations of higher frequencies are conveyed to the crystal with a relatively larger amplitude than those of low frequencies. This linear distortion, together with that originating from the cutting characteristic, is of such a nature that in the reproduction it can quite easily be neutralized by means of a simple R-C filter.

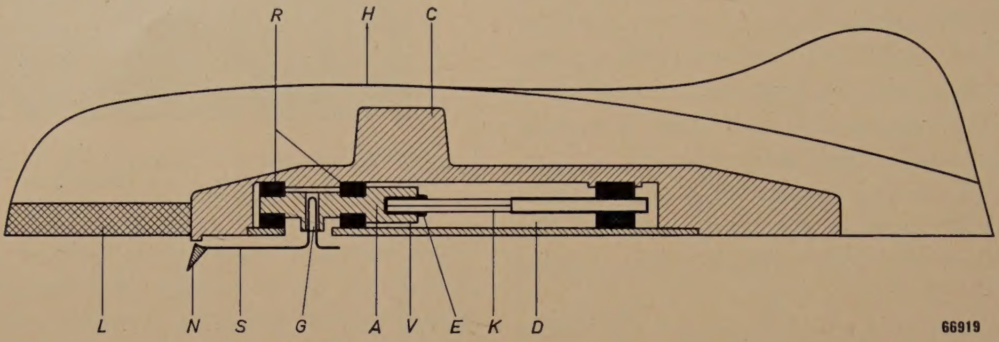


Fig. 11. The pick-up head with its most important parts. All that can be seen of it from the outside is the streamlined cover *H*. In the front of the head is a piece of lead *L* serving to increase the moment of inertia about the vertical axis. The moving parts of the pick-up head are mounted in the chassis *C*. The small rubber rings *R* serve as bearings for the armature *A*. Fitted in the aperture *G* of the armature is the drag spring *S* with sapphire needle *N*. The movements of the armature are transmitted via the fork *V* to the piezo-electric crystal *K*. Between the crystal and the fork is an elastic layer *E* which, owing to its frequency-dependent rigidity, ensures distortion-free sound reproduction. The space *D* around the crystal is filled with a gel protecting the crystal and damping parasitic vibrations.

Thanks to the presence of the elastic layer the torsional angles made by the crystal at low frequencies are relatively small, but owing to the great sensitivity of the crystal they are still large enough to induce the voltages necessary for good reproduction.

The gel with which the space around the crystal is filled has a two-fold function. In the first place it serves for damping parasitic vibrations likely to arise from resonance effects of various parts. In the second place it affords protection of the crystal against harmful atmospheric influences.

This need of protection has been further provided for in the packing of the pick-up, which is such as to ensure that the crystal will not suffer from long storage in any climate. The packing consists of an air-tight sealed metal tube containing a chemical which maintains the correct moisture content.

It has already been seen how the relatively large accelerations inevitably resulting from the movements of the armature essential for the reproduction have been kept within bounds. When the record is being played however, in addition to these accelerations and the static needle pressure already mentioned, there are also vertical forces of inertia as a result of the fact that neither gramophone records nor turntables are absolutely flat. Notwithstanding the much lower frequencies of the resulting movements, the forces that are brought into play may be considerable, on account of the attendant movement of the mass of the whole pick-up

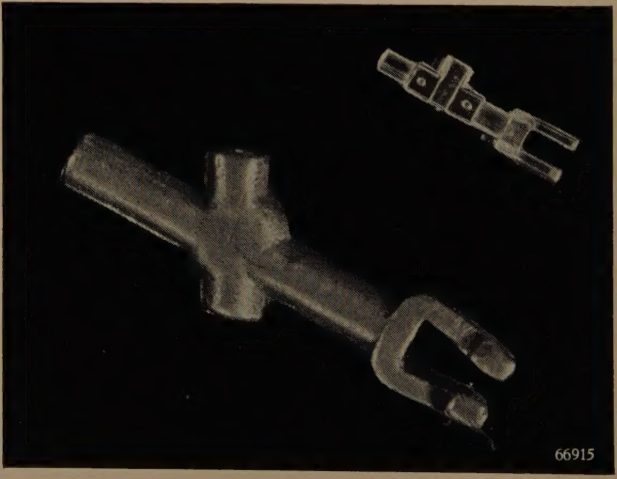


Fig. 12. A new armature (right) and an old one (left). The enlargement of the photograph is the same as that of fig. 10.

head plus the arm. Naturally the problems presented by these forces also become more serious when a sharper needle is used. It will now be explained how these problems have been solved in construction of the pick-up arm.

The pick-up arm

The pick-up arm is made of a very light thermoplastic material. The arm proper, which is remarkably slender, is set at a certain angle to the streamlined head carrying the crystal (see fig. 8). In the front of the head is a small piece of lead. At the other end of the arm are the bearings for the horizontal and vertical movements. A striking feature is the absence of a ball bearing for mounting the arm and the absence of any counterweight for reducing the static needle pressure.

Any such counterweight is superfluous here because the weight of the whole is so small that the correct needle pressure is obtained automatically. In fact the use of a counterweight is a most unsatisfactory method of reducing needle pressure, because then the forces of inertia arising from the periodical accelerations in the vertical direction are increased. Denoting the frequency at which these accelerations take place by $\Omega/2\pi$, the amplitude of the torque that the pick-up arm brings to bear upon the record may be expressed as $K \sim I_{\text{hor}}\Omega^2$, where I_{hor} represents the moment of inertia about the horizontal axis of rotation. Thus the force exercised upon the record is proportional to I_{hor}. From this it is seen that although a counterweight would reduce the static needle pressure, owing to the increase of I_{hor} it would amplify the variations of the dynamic pressure due to the wobbling of the disc. At the moments when this dynamic pressure is increased the record suffers more wear, and half a cycle later the needle pressure may be so small as to cause the needle to be forced upward along the wall of the groove by the horizontal accelerations, thus giving rise to distortion and background noise. In the worst case, if I_{hor} becomes too great, it may even happen that the needle jumps out of the groove.

After this exposition, from which it follows that the moment of inertia about the horizontal axis should be kept as small as possible, it may seem surprising that a weight has been introduced in the head of the pick-up, thereby increasing the moment of inertia I_{hor}. To understand properly the function of this weight it is necessary to consider the moment of inertia about the vertical axis.

The needle, together with the parts rigidly coupled with it in the horizontal direction (total mass m), follows precisely the modulations of the groove. The mass m is thereby displaced with respect to a mass M, the centre of gravity of which projected onto the plane of the gramophone record is, in the ideal case, always at a certain distance from the centre line of the unmodulated groove (see *fig. 13*).

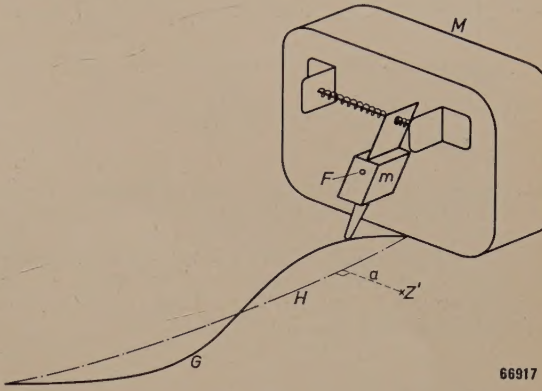


Fig. 13. Schematic representation of a pick-up. A small mass *m* is connected to a large mass *M* via a rotational movement about an axis *F*. Provided it is sufficiently small, the mass *m*, being rigidly coupled to the needle, follows precisely the modulations of the groove *G*, thereby describing a movement with respect to *M*, the centre of gravity of which when projected onto the plane of the record (Z') is at a constant distance *a* from the centre line *H* of the unmodulated groove.

In order to approximate this ideal case as closely as possible, the pick-up arm to which M is fixed must have the largest possible moment of inertia about the vertical axis, I_{vert}, for it is only in that case that the pick-up arm will be prevented from following the rapid deflections of the needle in the horizontal direction.

The object of the lead weight, therefore, is to give the pick-up arm a sufficiently large moment of inertia about the vertical axis. Furthermore, being placed in front of the needle, it causes the bearing pressure of the arm to be so small that a complicated and thus expensive ball bearing can be dispensed with. It may seem strange that what is done here is just the reverse of what was to be expected from the arguments previously expounded: the bearing pressure is reduced at the cost of increased needle pressure. But the fact that the weight is placed in the front end of the head results in a minimum increase in needle pressure. The explanation of this is illustrated in *fig. 14*.

The two requirements of small I_{hor} and large I_{vert} are naturally opposed to each other, but the construction indicated proves to be a compromise which in practice satisfies the most stringent demands.

The extent to which, for instance, it has been possible to reduce I_{hor} may be judged from the following simple but instructive test: when a cigarette is placed between the turntable of the gramophone and a record thus causing an abnormal wobbling of the record, it will be seen that the quality of the reproduction is scarcely affected at all.

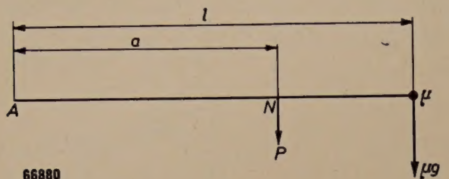


Fig. 14. Increase of the moment of inertia about the vertical axis A. N = needle, μ = the mass applied in order to increase I_{vert}. Assuming that for good sound reproduction a certain increase I_{v} of the moment of inertia about the vertical axis is needed, then $I_{\text{v}} = \mu l^2$. The increase of the needle pressure due to the mass μ is $P = \mu l/a = I_{\text{v}}/al$. Thus, with a given I_{v}, P decreases with increasing l.

Another point clearly but simply demonstrating the advantages of the new type of pick-up is connected with the needle talk, which is mainly a result of the vertical movements of the needle and the mass rigidly coupled to it in those directions. Owing to the great reduction of the mass this needle talk has been so minimized that there is no need to enclose the turnable and the pick-up in a cabinet, as is the case with other types of pick-up (see fig. 8).

Finally it is to be mentioned that a record of reasonable quality suffers so little wear that even after it has been played a hundred times there is hardly any noticeable difference in the reproduction. Thus it may be claimed that owing to the extreme care taken in the construction of the pick-up gramophone records have been given an almost unlimited life.

Summarizing, it is seen that thanks to the measures described it has been possible to combine three important features, viz: a longer playing time, better reproduction and very little record wear.

Summary. The playing time of a gramophone record is limited by the width of the groove, the width of the wall between adjacent grooves and the minimum groove speed (linear velocity of the innermost turn). The groove width can be reduced only if a needle is used which has a smaller rounding-off radius of the tip, and this is impossible unless the weight of the pick-up is reduced. The width of wall is determined by the desired amplitude in recording and therefore, due to the signal-to-noise ratio, by the granulation of the record material. The width of the wall can therefore be reduced if a material of a finer structure is used. Owing to tracing distortion and pinch effect the minimum groove speed can be reduced only if all parts of the pick-up are simultaneously so changed that none of the pressures arising during the playing of the record exceeds the limit of elasticity of the record material. The pick-up of the new Philips record players satisfies all these requirements, so that long-playing records ($22^1/_2$ minutes for a 12″ cm record) can be played with an extremely small amount of distortion. The new record players also offer the possibility of playing records at different speeds, so that both ordinary and long-playing (microgroove) records can be played on the same instrument.